

Review

Not peer-reviewed version

Review of Health Knowledge Graphs for Clinical Decision Support: Construction, LLM-Grounded Reasoning, and Applications

[Subhankar Chatteraj](#)^{*}, Karuna Pande Joshi, Sawon Pratiher

Posted Date: 16 October 2026

doi: 10.20944/preprints202610.1625.v1

Keywords: Health Knowledge graph; large language model; agentic framework; deep learning; health care; EHR; clinical decision and application



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Review of Health Knowledge Graphs for Clinical Decision Support: Construction, LLM-Grounded Reasoning, and Applications

Subhankar Chattoraj ^{1,*}, Karuna Pande Joshi ¹ and Sawon Pratiher ²

¹ University of Maryland Baltimore County

² Indian Institute of Technology Kharagpur, Kharagpur, 721302, India

* Correspondence: chattorajsubhankar@gmail.com

Abstract

Health Knowledge Graphs (Health-KGs) structure heterogeneous medical information from Electronic Health Records (EHRs), the biomedical literature, and clinical guidelines into machine-reasonable form, and have become a foundation for clinical decision support. This work reviews the construction and application of Health-KGs following the PRISMA 2020 guidelines, organizing the literature along three axes representation, extraction, and integration and situating it against prior biomedical-KG reviews. We examine knowledge-representation learning and graph-embedding models, named-entity recognition, disambiguation, and linking, and cross-source integration, together with the principal EHR datasets and clinical language models that underpin them. A distinguishing focus is the shift toward large language models (LLMs): we synthesize LLM-driven entity and relation extraction, retrieval and graph-augmented clinical reasoning, and the factuality and hallucination constraints that govern their safe use an axis not addressed by prior EHR-focused reviews. We further reviewed applications spanning drug development and clinical trials, diagnosis and prediction, treatment recommendation, question answering, and health-policy compliance. Finally, we identify the central open challenges: the absence of a standardized, openly governed Health-KG benchmark; heterogeneous and temporally evolving clinical data; multimodal fusion; and privacy-preserving, reproducible construction. This review is intended as a roadmap for data scientists, medical-informatics and AI researchers, and healthcare practitioners.

Keywords: Health Knowledge graph; large language model; agentic framework; deep learning; health care; EHR; clinical decision and application

1. Introduction

In numerous healthcare applications that contain multi-modal data, graph-based structures are suitable to represent the information in a structured manner. The data structure capturing the relationship between entities and attributes is termed as Knowledge Graph (KG) [1,2]. The information necessary to construct the KG is collected from websites, domain experts, and knowledge bases that are stored as unidirectional (one-way) or bidirectional (two-way) relationships, depending on the directionality of the association between the connected entities [1]. In recent years, KG has been applied to the multidisciplinary domain from the Google search engine to health care [3]. The application of KG models has been studied in a wide range of applications from basic science research [4–6], pharmaceutical research [7,8], clinical decision support [9,10], question answering and recommendation systems [11–15]. Post-digitization of healthcare systems, the longitudinal and comprehensive data collected during the delivery of patient care is created and encoded as Electronic Health Record (EHR) data. The National Academy of Medicine defines the core functionality of EHR as health information and data, decision support, electronic communication and connectivity, patient support administrative

processes and reporting, and population health management [16]. Meaningful use of EHR has early evidence in better healthcare safety and quality benefits [15,17].

Health-KGs are constructed from intricate medical systems such as single or multiple integrated EHR datasets, medical literature, clinical guidelines, and patient-generated information where the Natural Language Processing (NLP) techniques are used to optimize entity and relation extraction methods [18,19]. The open medical EHR databases such as 'Medical Information Mart for Intensive Care' III (MIMIC III-IV) [20,21], and medical terminology for diagnoses, symptoms and procedures for claims processing such as Unified Medical Language System (UMLS) [22], Systematized Nomenclature Of Medicine (SNOMED-CT) [23], ICD-10 [24], and drug database such as DrugBank [25], PharmGKB [26] and KEGG [27] are also useful for Health-KGs construction.

Recent advancements in Machine and Deep Learning (DL) have demonstrated potential for resource allocation via task prediction, patient diagnosis [28–31], length-of-stay [32], and mortality estimation [33–35]. Health-KGs have been widely adopted to improve the data insights and complement EHR modelling [36–38]. Pre-trained large language models (LLM) such as BERT [39] and GPT [40] have been implemented for knowledge extraction in Health-KGs construction. The utilization of pre-trained LLM has allowed more comprehensive and accurate medical ontologies and taxonomies, further enabling efficient and effective integration of heterogeneous medical data from multiple resources [41–43]. However, there are significant challenges in Health-KGs construction due to their diverse and scattered nature, making integration and effective analysis quite complex [44]. This heterogeneity can also result in incomplete or conflicting data representations within Health-KGs, thereby limiting their utility for various healthcare applications [45–47]. Furthermore, the prevailing utilization of domain-specific knowledge graphs may lead to constrained coverage and granularity of knowledge encompassing various medical domains. This limitation hampers the ability to identify correlations and connections among medical concepts spanning multiple areas, thereby impeding comprehensive healthcare insights. This paper presents the results of our comprehensive research analysis on constructing Health-KGs from EHR datasets, the present state-of-the-art from most recent and promising research direction in Health-KGs leveraging LLM to personalized KGs, challenges, and its applications. An extensive literature analysis concerning Health-KGs was carried out at three levels: representation, extraction, and population or integration. The comprehensive summary techniques involved in constructing Health-KGs from scratch and integration are explored. Additionally, the model-based and model-free Health-KGs are also highlighted. Furthermore, an in-depth overview of the diverse applications of Health-KGs in drug development, clinical trial diagnosis, predictions, and treatment recommendations are also explored. It was observed that there is a noticeable gap in research within three critical areas: (i) managing complex real-world patient data in Health-KGs, (ii) leveraging semantic features for knowledge alignment, and (iii) addressing the dynamic nature of healthcare entities and relationships. This study on Health-KGs for healthcare advancement and meaningful application is a practical roadmap and resource for the healthcare community. Our primary research contributions are as follows:

- An extensive study of existing literature is explored to understand the present State-of-the-Art and identify the gaps and challenges in Health-KGs.
- The various methodologies, algorithms, and techniques used in Health-KGs construction, representation learning, reasoning, benchmarking, evaluation, and applications are explored in detail.
- A comprehensive taxonomy of the Health-KGs and its various applications is constructed in Figure 2.
- A dedicated synthesis of LLM-driven Health-KG construction and KG-grounded clinical reasoning (entity/relation extraction, retrieval- and graph-augmented generation, and factuality), an axis not addressed by prior EHR-focused reviews.

The results of this study will benefit a wide range of audiences, from data scientists, medical informatics, artificial intelligence researchers, and healthcare practitioners. The rest of the paper is structured as

follows: Section II describes the search methodology; Section III situates this reviews against prior reviews; Section IV covers EHR data; Section V details Health-KG construction; Section VI examines the role of large language models; Section VII reviews applications; Section VIII discusses evaluation, open challenges, and future directions; and Section IX concludes.

2. Search Methodology

We conducted the review following the PRISMA 2020 reporting guidelines [48]. We searched PubMed, Scopus, IEEE Xplore, and the ACM Digital Library for literature published between January 2013 and December 2025, using combinations of the terms "electronic health record," "EHR," "knowledge graph," "medical/clinical knowledge graph," and the task qualifiers "construction," "representation," "extraction," "integration/population," and "large language model." The exact query strings per database are given in Table 1. We also hand-searched the proceedings of ACM SIGKDD, IEEE EMBC/EMBS, IJCAI, and AAAI and the archives of high-impact venues (IEEE Access, Knowledge-Based Systems, Expert Systems with Applications, Journal of Web Semantics, Journal of Biomedical Informatics).

Inclusion criteria: peer-reviewed studies that (i) construct, represent, integrate, or reason over a knowledge graph using EHR or closely related clinical data, or (ii) propose methods or benchmarks directly applicable to Health-KG construction or application. Exclusion criteria: non-English articles, non-clinical knowledge graphs, opinion pieces without methodological contribution, and duplicates across databases. After de-duplication, records were screened by title and abstract, then assessed in full text against the criteria above. Figure 1 reports the counts at each stage of the selection process.

Table 1. Search Strings per Database.

Database	Query (representative)
PubMed	("electronic health record"[tiab] OR EHR[tiab]) AND ("knowledge graph"[tiab] OR "knowledge graphs"[tiab]) AND (construction OR representation OR extraction OR integration OR "large language model")
Scopus	TITLE-ABS-KEY(("electronic health record" OR ehr) AND ("knowledge graph") AND (construction OR representation OR extraction OR integration OR llm))
IEEE Xplore	("All Metadata":electronic health record) AND ("All Metadata":knowledge graph) AND (construction OR representation OR extraction OR integration)
ACM DL	[Abstract: "electronic health record"] AND [Abstract: "knowledge graph"] AND [Abstract: construction OR extraction OR integration]

Figure 1: PRISMA Flow Diagram for Systematic Review

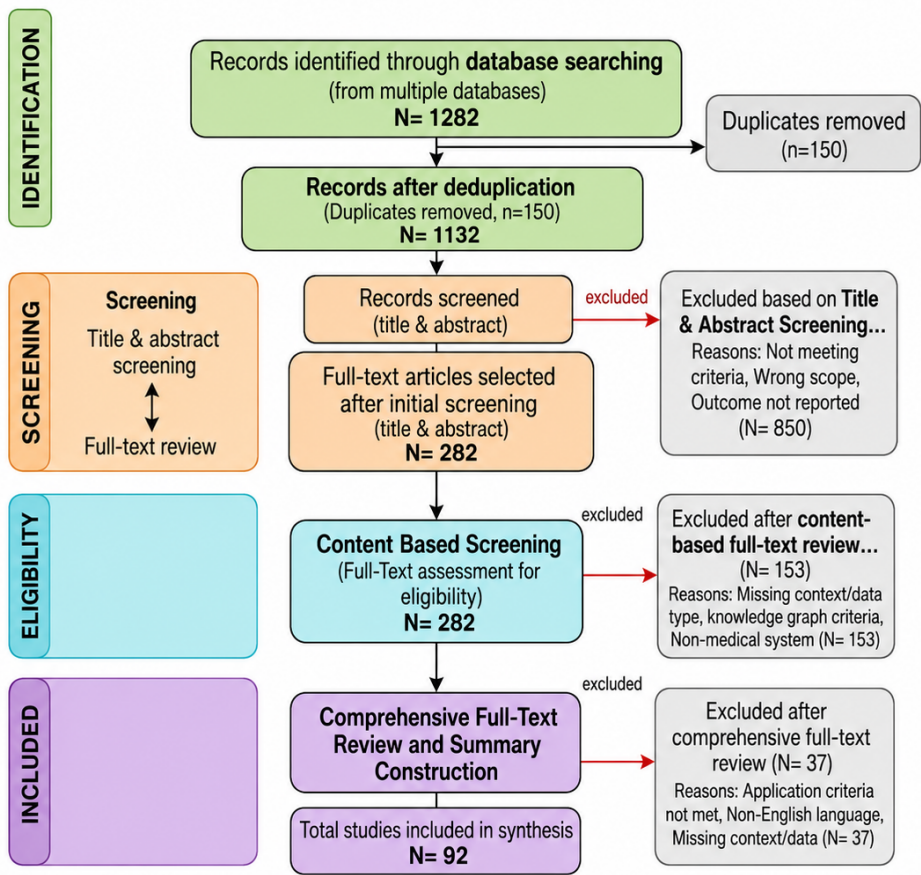


Figure 1. Article selection and search methodology.

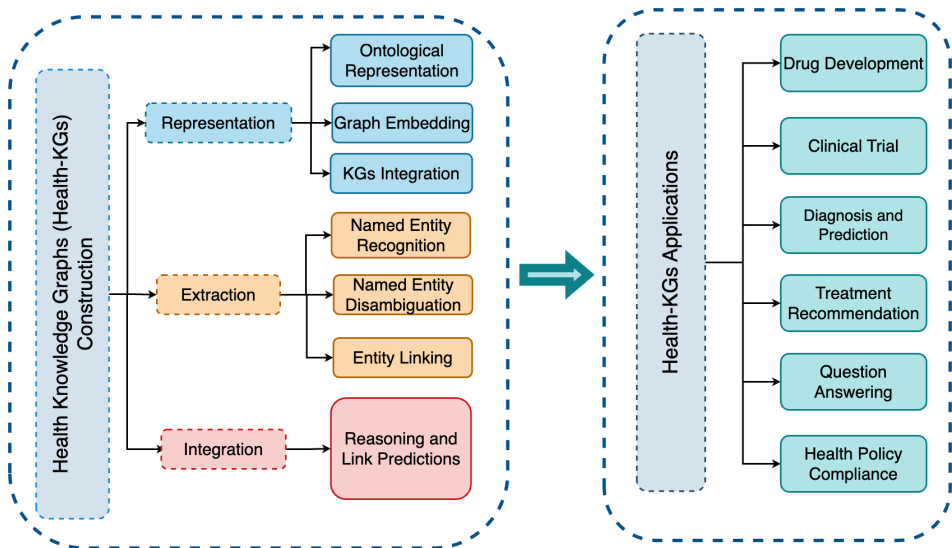


Figure 2. The Comprehensive Taxonomy for Health-KGs Construction & Application.

3. Related Reviews

3.1. Related Reviews and Positioning

Knowledge graphs in the biomedical domain have been reviewed from several angles. Nicholson and Greene [1] surveyed the construction of biomedical knowledge graphs and their applications broadly across molecular biology. Ji et al. [2] provided a general (non-clinical) survey of knowledge-

graph representation, acquisition, and applications. Abu-Salih [3] reviewed domain-specific knowledge graphs across multiple fields without an EHR focus, and Silva et al. [49] surveyed ontologies and knowledge graphs specifically for oncology research. More recently, Pan et al. [50] proposed a roadmap unifying large language models and knowledge graphs at a general (non-medical) level. None of these jointly addresses EHR-centric construction along the representation–extraction–integration axes, the recent shift toward LLM-based construction and reasoning, and a unified clinical application taxonomy. Table 2 summarizes this positioning.

Table 2. Scope of this Review Versus Prior Reviews. (✓ = covered, ◐ = partial, X = not covered).

Reviews	Year	EHR-centric	Construction	Extraction	Integration	LLM coverage	Applications	Challenges
Nicholson & Greene [1]	2020	◐	✓	✓	◐	X	✓	◐
Ji et al. [2]	2021	X	✓	✓	◐	X	◐	◐
Abu-Salih [3]	2021	X	✓	◐	◐	X	◐	◐
Silva et al. [49]	2022	◐	◐	◐	X	X	◐ (oncology)	◐
Pan et al. [50]	2023	X	◐	◐	◐	✓	◐	◐
This work	-	✓	✓	✓	✓	✓	✓	✓

4. EHR Data: Sources, Standard, Characteristics, and Complexity

The patient’s symptoms, test results, diagnosis, treatment, medication, and follow-up visits are stored in the hospital system as EHR data. The EHR data are comprised of several controlled vocabularies and classification schemas, such as procedure codes (Current Procedural Terminology, statistical Classification of Diseases and Related Health Problems) [51,52], and laboratory observations (such as the Logical Observation Identifiers Names and Codes) [53] to record the patient health-related information. EHR data can be structured in the form of lab reports or unstructured in the form of progress notes. This section discusses the sources of publicly available EHR data, the common standard of EHR management, and the challenges of creating Health-KGs from EHR data.

4.1. EHR Data Sources

The primary source of EHR data is stored within hospitals. The most significant challenge is accessing the EHR data as all organizations generally have strict human research organizational policies due to the presence of Protected Health Information (PHI). However, there are some large-scale publicly available EHR datasets for research purposes. Table 3 summarizes the principal publicly available EHR datasets, their approximate scale, modalities, and access terms. MIMIC-III [20] is a comprehensive dataset of EHR for patients admitted to tertiary care hospitals in the US in critical care units. The database comprises billing (CPT, Diagnosis-Related Group (DRG) codes, and International Classification of Diseases (ICD) codes), demographic details, dictionary, interventions, laboratory, medications, notes, physiologic, and reports (vital sign assessment). MIMIC-IV [21] database constructed from EHR records from Beth Israel Deaconess Medical Center (BIDMC) and Massachusetts Institute of Technology (MIT) comprising patient diagnoses, free-text clinical notes, orders, measurements, treatments, and procedures. The national NLP clinical challenges (n2c2) [54] are also available, subject to the data user agreement. The field’s heavy reliance on the MIMIC family means most reported Health-KG results are ICU-centric and single-region; broader external validation on multi-center resources (eICU-CRD) and non-ICU populations remains scarce, which limits claims of generalizability.

4.2. Healthcare Ontologies and EHR Standard

The EHR standards such as Health Level Seven (HL7) [55], openEHR [56], and CEN/ISO EN13606 [57] cover the multiple aspects of clinical records management. These standards cover integrating medical devices, exchanging patient data between healthcare institutions, and interfaces with clinical decision support systems. HL7 supports health service evaluation, delivery, and clinical practice and management. The structure and semantics of the EHRs are standardized using the Clinical Document Architecture (CDA) document. The HL7 RIM [58], HL7 DT is used to define the machine derivable

CDA and using the SNOMED CT [23] it facilitates clinical documents semantic representation and interoperable with other standards. The OWL/RDF automatic rule-based inferencing is explored for HL7 RIM and HL7 DT representation. Furthermore, with the advancement of Web technologies, HL7 has also been extended to Fast Healthcare Interoperability Resources (FHIR) [59].

The primary aim of the openEHR and CEN/ISO EN13606 is to standardize the management, storage, retrieval, and exchange of health data in EHRs. The European Committee for Standardization (CEN) stated the norm for exchanging data between EHR systems known as Electronic Health Record Communication (EN13606). However, a subset of openEHR can be viewed as a stand-alone standard. In the past, the reference model and data types and structures of the openEHR OWL ontology following the official openEHR were developed [60]. Further, an ontology was formed recently to provide semantic interoperability between EN13606 and openEHR standards [61].

Table 3. Publicly available EHR datasets for Health-KG research.

Dataset	Scale (approx.)	Modalities	Access
MIMIC-III [20]	~46k patients, ~58k ICU admissions (2001–2012)	Structured (codes, labs, vitals) + free-text notes	Credentialed (PhysioNet DUA)
MIMIC-IV [21]	~300k patients, ~430k admissions (2008–2019)	Structured + notes; modular tables	Credentialed (PhysioNet DUA)
eICU-CRD [62]	~200k ICU stays, 208 hospitals	Structured, multi-center	Credentialed (PhysioNet DUA)
n2c2 / i2b2 [54]	Task-specific identified de-note corpora	Unstructured clinical text	Data-use agreement
PLAGH (private) [63,64]	Institutional cohort	Structured + notes	Restricted/private

4.3. Multimodal EHR Data

The patient health records include demographics, laboratory test results, vital symptoms, histopathological test results, radiological test reports with different imaging modalities, and current and follow-up diagnoses. In a recent approach to model development, researchers have considered including most of the modalities despite the complexity of the nature of the data. Integrating multi-modal data in Health-KGs construction generally involves two primary phases;

- Using representation learning to properly represent multi-modal data from structured knowledge and language from many sources to overcome structural disparities [65,66]
- Depending on the data representation structure, adapting multi-modal learning for improved reasoning, classification, and prediction [67].

Recently, DL-based methods have been applied to tabular domain data, where a multi-modal transformer framework based on self- and semi-supervised learning frameworks was proposed to explore value imputation and data augmentation was studied using real-world MIMIC-III data [68]. Given the complexity of raw real-world data, the underlying explicit feature interactions from EHR data were modelled using Deep Cross Network (DCN) [69] and the novel pipeline to derive usable hourly time-series data from the raw MIMIC-III data [70]. In a recent article, ‘FIDDLE’, a systematic pre-processing technique for processing EHR data, was proposed [71], utilizing for medication [72] and diagnosis predictions [73]. The existing diagnosis predictions using a learning-based algorithm were improved by leveraging the multi-modal nature of the EHR data [74,75]. A multi-modal attention neural network trained to combine MIMIC-III clinical notes and codes was proposed [76].

4.4. Heterogeneous Temporal Data in EHR

The heterogeneous medical data such as patient ID, progress and current diagnosis, lab reports, imaging results, and medications are stored in EHR [77]. More importantly, constructing and utilizing Health-KGs to predict clinical decision support tasks incorporating clinical data collected over time with cross-departmental information provided enhanced, in-depth, superior longitudinal analyses and patient trajectories of clinical events. However, the modelling and integration of heterogeneous data is challenging due to the complex nature of the data. Besides, temporal KGs are proposed by integrating time stamps in the model, which better captures the underlying complex temporal dynamics of the entities and relationships in EHR data [78].

4.5. Semantic Complexity of EHR

Improved contextual knowledge utilization for improved interoperability, reduced duplication, and missing entity identification are made possible by the semantic knowledge representation of the combined and linked entities. The main challenges faced by the researcher in constructing Health-KGs is the complex semantic patterns, which frequently manifest as terms and medical disorders in clinical records. In a recent work, the conceptual graphs were integrated with health data. To separate the contextual information from the complicated semantic textual data, they employed inference mechanisms and textual pattern mining for knowledge retrieval [79]. The absence of prior subject expertise was the reason for the unimpressive outcomes. Removing repetitive semantic structures from EHR was proposed using syntactic trees from sentences to search similar sub-trees and labelled sentences by employing word embedding algorithms [80]. Learning medical knowledge graphs from EHR data is more complicated than learning from structured sources such as textbooks and articles. However, there are a lot of associated relationships that a KG created from real patient information infers. The benefit of creating a Health-KG from an EHR is the minimum post-processing requirements.

5. Health-KGs Construction: Representation, Extraction, and Integration

To understand the significant features of EHR using Health-KGs, we analyzed the existing approaches used to construct Health-KGs. Figure 2 illustrates the comprehensive taxonomy for Health-KGs in four primary categories: representation, extraction, integration, and applications. The primary step of Health-KGs construction is identifying data from sources such as open databases, scientific publications, and real-world EHR [81]. The data, once identified, are primarily available in the form of structured data (such as demographic information, vital signs, laboratory test results, and diagnostic codes) and unstructured data (such as clinical notes, medical images, and radiology reports) [82]. The subsequent sections below explain more comprehensive details on the construction of Health-KGs.

5.1. EHR Representation and Learning

The heterogeneity, temporality, and knowledge intensity are the primary challenges in representing EHR data [83]. The most common approach to representing data is in the form of two entries and a relation connecting these entries, i.e., triplet format. The triplets are extended to a quadruple format, which includes additional disease or patient-related properties [84]. The representative sample of triplet and quadruple format is shown in Figure 3. The main focus on existing Knowledge Representation Learning (KRL) is to learn the semantic relations, relational triplets (represented as head entity, relation, tail entity), and attribute triples (delineated as entity, attribute, attribute value) [85].

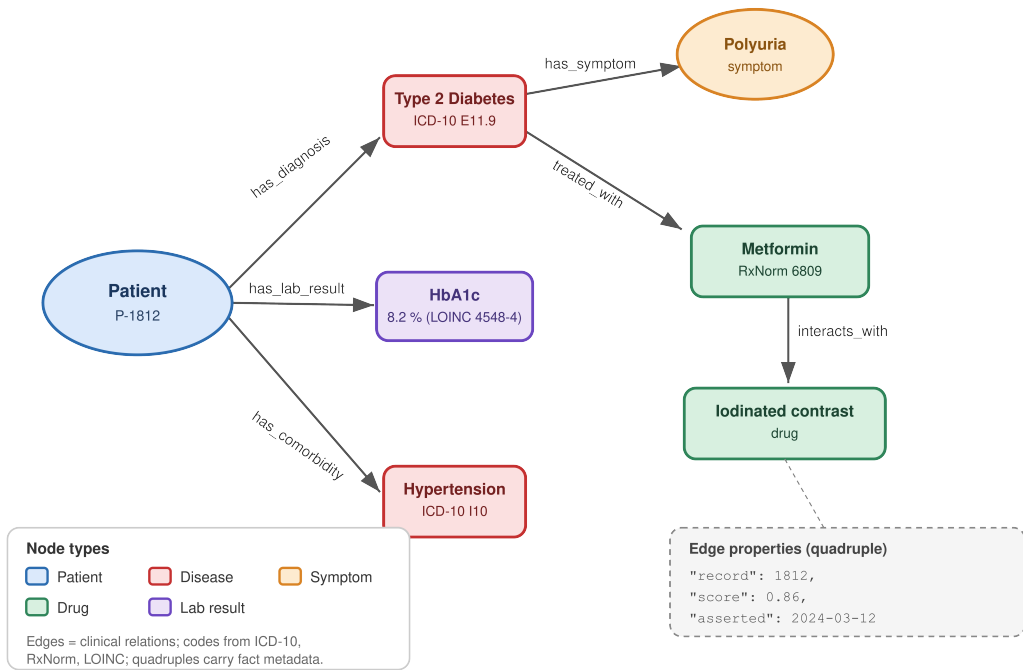


Figure 3. A representative Health-KG subgraph constructed from a single EHR encounter, illustrating typed entities, normalized ontology codes, clinical relations, and quadruple-level fact properties.

The existing KRL approaches are divided into two categories. The first involves modelling the structural information among entities, while the other involves using texts or semantic information retrieved from triples in KGs. The proposed Trans series models are computationally easier to train, representing the anterior approach yet work in one-to-one relations [86]. In the TransH model, the head and tail nodes are projected onto the same vector in the relation plane, which addresses the 1-to-n and n-to-1 problem between entities and relations [87]. The TransE and TransH models project the head and tails in the same vector space, which was addressed using the TransR model [88]. The other improved models in the Trans series models included PTransE [88], TransD [88], and TransSparse [88]. The different text semantic pertinences between different entities can be modelled using the neural tensor network model [89]. The targeted text is extracted via embedding of target entities through multiple convolutional filters using end-to-end fully connected convolutional neural network (CNN) proposed in the ConMask model [90]. The further optimization in the form of the conditional probability of the objective function proposed in KR-EAR methods demonstrated superior performance in entity, relation, and attribute prediction tasks, respectively [91]. The triplet prediction and attribute value regression were modelled using a multi-task neural network approach [92], while the Description-Embodied Knowledge Representation Learning (DKRL) performs the representation learning of similar task via a weighted combination of the TransE model and hybrid model, which was a combination of Continuous Bag of Words (CBOW) model and deep CNN [93]. The comprehensive structural and semantic relational and attribute triples information was utilized using KG via Attribute-embodied Knowledge Representation Learning (AKRL) [84].

5.2. Ontological Representation

The sharing of the EHR across different healthcare providers securely and efficiently is an ongoing challenge. Integrating medical ontology models for constructing Health-KGs from heterogeneous EHR data across various sources assists in achieving interoperability and semantic visualization, which is often beneficial for clinical research.

Several works have been proposed for creating medical ontologies. The properties of tumorigenesis and cancer cell description are determined in OncoCL ontology [94], which has been built by using Cell Ontology (CL), UBERON, BRENDA Tissue Ontology (BTO), Pathway Ontology, Phenotype And Trait Ontology (PATO), Coding Process Ontology (CPO), and Sequence Ontology (SO). The drug target prediction is determined from the CRC ontology [95] by reusing the PharmGKB ontology. The POCS ontology [96] was developed, reusing the Friend Of A Friend (FOAF) ontology for cancer survivor engagement. The HemOnc ontology was constructed to systematize hierarchic cancer treatment [97]. Besides, the survival outcome measures of clinical trials in oncology were proposed in TOCSOC ontology, which was constructed using Experimental factor ontology (EFO), CCTOO, Interlinking Ontology for Biological Concepts (IOBC), and National Cancer Institute Thesaurus (NCIT) [98]. Newer terminology-focused and Semantic-focused application ontologies are available in [49].

5.3. Graph Embedding

Graph Embedding (GE) converts the KG into a low-dimensional vector to represent the data. GE preserves the semantic information by computing the entity similarity in the form of distributed embeddings as Knowledge Graph Embedding (KGE) [99,100]. The primary aim of KGE is to represent the non-semantically connected entities and relations in similar vectors. In general, the neural network is trained over a KG to learn the KG via three primary steps:

- Encoding entities into dispersed points within the semantic space, with relations encoded as vectors
- Employing a scoring function or model-specific function to amalgamate information derived from a triplet
- Implementing an optimization procedure guided by the loss function. This optimization aims to define and minimize the objective during the knowledge graph embedding, ensuring effective representation and learning from the encoded entities and relationships [101].

The categorization of embedding techniques typically encompasses translational distance (TD), semantic matching (SM), neural network-based models (NNM). TD utilize distance-based score functions, while SM employ similarity-based score functions. These mechanisms calculate the distances or similarities between entities and relationships, forming the basis for generating embeddings in knowledge representation [100,102]. Table 4 compares the principal KGE models by category, relation patterns handled, relative complexity, and key limitations.

In a recent work, the projection of a single entity into several projection vectors in various relations was made possible by a relation-dependent hyperplane constructed via mapping the subject and object embedding vectors [103] while the other work, in contrast to hyperplane, it was focused on relation-specific space [104]. The probabilistic model learned from Health-KGs was proposed to predict the missing link [105]. Assessing the formed triplets' validity involves using semantic or distance-based scoring functions. Subsequently, to enhance the model knowledge, the representations of facts are encoded with a suitable neural model, incorporating additional information such as entity types and attributes descriptions, & relational paths [2]. Semantic matching models that capture the inherent semantics of triples include RotatE [106], DistMult [107], and RESCAL [108], ComplEx [109]. The ConKB and ConvE [110,111] utilize convolutional neural networks (CNN) based architecture to feature analysis of triples. The extraction of semantically relevant terms using word embeddings trained on EHR and medical literature has been prevalent. The Multiple BERT version SciBERT [112,113], ClinicalBERT [114], and BioBERT [115] are proposed in the clinical domain. Integrating KGE into clinical KGs representation has demonstrated significant success in the clinical pharmacy domain in applications like DL-based drug discovery [116,117] and prediction of drug-drug interactions [118]. Node embedding plays a crucial role in testing various link prediction methods within biomedical KG [119]. For instance, the node2vec [120] algorithm generates node embeddings from a KG centred around drug reactions [121]. Case studies conducted on COVID-19 and liver injury drugs showcased the algorithm's effectiveness [122]. Furthermore, utilizing graph-based representation to learn the

concept embeddings from SNOMED-CT in healthcare analytics yielded promising link and patient state prediction [123]. The joint embedding-based representation model, which incorporates descriptions of semantic medical entities and triples, was developed to measure patient similarity. The joint embedding-based representation model was developed using the Siamese CNN model utilized MIMIC-III dataset [20], yielding superior results in patient clustering [124]. Using GE to construct specific disease knowledge graphs in multi-modal reasoning has shown promising outcomes. However, KGE is in the early development stage to provide universal and new reliable knowledge in the disease-specific field [4]. In Health-KGs the choice of embedding is governed less by raw link-prediction accuracy than by the relational structure of clinical data, where many-to-many and antisymmetric relations (e.g., drug–interacts-with–drug, disease–progresses-to–disease) dominate; this favors RotatE and ComplEx over the original TransE, while CNN-based models such as ConvE trade interpretability for expressiveness.

Table 4. Comparative Analysis of Knowledge Graph Embedding Models for Healthcare Knowledge Graphs/

Model	Category	Relation Patterns Handled	Relative Complexity	Key Limitation
TransE [86]	Translational distance	1-to-1	Low	Fails on 1-to- <i>N</i> , <i>N</i> -to-1, <i>N</i> -to- <i>N</i> , symmetric relations
TransH [87]	Translational distance	+ 1-to- <i>N</i> / <i>N</i> -to-1 (via hyperplanes)	Low–moderate	Still limited for <i>N</i> -to- <i>N</i> , composition
TransR [88]	Translational distance	Complex relations (separate entity/relation space)	High (per-relation matrices)	Many parameters; overfitting risk
TransD [88]	Translational distance	Complex relations (dynamic mapping)	Moderate	Fewer params than TransR; weaker on composition
RotatE [106]	Semantic / rotation	Symmetry, antisymmetry, inversion, composition	Moderate	Complex-space tuning sensitivity
DistMult [107]	Semantic matching (bilinear)	Symmetric only	Low	Cannot model antisymmetric relations
ComplEx [109]	Semantic matching (complex bilinear)	Symmetric + antisymmetric	Moderate	Less interpretable embeddings
RESCAL [108]	Semantic matching (full bilinear)	Expressive (full tensor)	High	Quadratic relation params; overfitting
ConvE [111]	Neural (2D CNN)	Rich interactions	Moderate	Less transparent; reshaping artifacts

5.4. Health-KGs Integration

The KG integration is crucial in constructing Health-KGs to avoid repetitive work and construction from scratch. Integrating or merging multiple Health-KGs are effectuated into single Health-KGs [125–127]. The dissimilar data format, terminologies, and schemas make the process of Health-KGs integration more challenging. In recent work, various methods have been proposed to address this challenge. The fusion algorithms like schema alignment [128,129], entity and conflict resolution [130–

[132], and ontology matching [133,134] provides robust integration and reconciliation to integrate different Health-KGs. In earlier proposed work, Health-KGs were integrated with DL-based methods to diagnose thyroid disease [135]. The integration of context retrieved from the large corpus of cancer pathology reports using KGE and external medical terminology ontologies was used to train a multi-task CNN for cancer phenotyping [136]. The partially overlapping multiple EHR institutional codes were integrated with the textual data using the multiview incomplete Health-KGs integration algorithm. The results demonstrated superior cross-institutional data harmonization and improved semantic accuracy related to COVID-19 EHR studies [137].

5.5. Extraction

In the context of knowledge extraction, Health-KGs construction involves the extraction of entities and the relationships between them. The extraction involves three main tasks [138]: i) Named Entity Recognition, which involves recognition of organizations, events, individuals, location and other entities from (un)structured or semi-structured data; ii) Named Entity Disambiguation, aimed at resolving the ambiguity of an inferred entity by mapping it to its real-world counterpart; and iii) Named Entity Linking, assigning a unique Internationalized Resource Identifier to the disambiguated entity.

5.5.1. Named Entity Recognition

Named Entity Recognition (NER) is a popular research domain in the scientific community. Most approaches have been domain-specific and limited to reviews, news, and health care. We categorize the literature into three main categories.

- The rule-based KG approaches, in this category, most studies have implemented hand-crafted rules-based approach [138,139]. This approach does not require Large-scale annotated data since they mostly rely on lexical resources. Additionally, due to the nature and domain-specific approach, very high precision is often achieved [140]. However, the most common disadvantages are the unavailability of lexicon resources, domain dependence, and the cost of expert resources for many languages.
- The second approach is learning-based approaches, which are categorized into supervised, semi-supervised, and unsupervised. Machine learning (ML) models like Support Vector Machines, Conditional Random Fields, Hidden Markov Models, and decision trees are trained on input samples and used as classifiers [141]. More end-to-end models via bootstrapped and unsupervised methods are also proposed, requiring minimal training or seed ontologies [142–144].
- The feature-inferring neural-network approaches is the most recent and last category in this domain [39,145–149]. In contrast to the rule- and learning-based methods, these approaches leverage ML but distinguish themselves by automatically deducing features through DL. Recent studies suggest that they surpass earlier methods [141,142,150]. Unlike the aforementioned approaches, these methods do not necessitate seeds, ontologies, or domain-specific lexicons, rendering them more domain-independent. Additionally, they capitalize on the precision of their inferred features. However, constructing the large-scale dataset is required to develop robust models using these approaches.

5.5.2. Named Entity Disambiguation

In conventional Named Entity Disambiguation (NED) research, the computation of similarity between a mention and its candidate entities often relies on manually crafted features [151,152]. These investigations can be categorized into independent (or individual), collective, and Neural Network based approaches.

- Independent methodologies employ semantic similarity techniques to arrange candidate entities based solely on their lexical similarity and/or observed co-occurrence with mentions [153,154]. Each mention undergoes autonomous disambiguation in these approaches, framing entity disam-

biguation as a ranking challenge that selects the entity with the highest confidence. Confidence levels are determined by amalgamating handcrafted features drawn from the mention's context, including surrounding words and textual portrayals of the candidate entity, often extracted from sources like Wikipedia. Various similarity metrics, such as BOW-based cosine similarity, have been suggested. Nevertheless, owing to the limited textual information inherent in handcrafted features, the performance decreases in intricate scenarios [152]. Moreover, a more complex model-based approach is required to accommodate better the intricate and interconnected nature inherent in natural language.

- Collective methodologies utilize semantic similarity considering that text generally revolves around the same subject [155]. Consequently, co-occurring entities are expected to frequently exhibit semantic relationships [156]. The associations between candidate entities are thus established using collective approaches, and the probabilistic graph method is utilized to build the mention-entity pairs. The Babelify [157], AGDISTIS [158], and TagME [159] leverage graph connectedness benefits from semantic relationships among disambiguation candidates outlined in a knowledge base. Most of them rely on BOW, which is computationally expensive.

Neural network-based models have been prevalent and have gained significant success. Recently, to reduce the time in NED recognition in Health-KGs applications, the Long short-term memory (LSTM)-conditional random field (CRF) model was utilized. The proposed model was a hybrid combination of rule-based and dictionary-based and achieved a significant recognition rate [160]. Besides, a similar hybrid model comprising verb and noun phrases for the relationship between classification was proposed [161]. The feature set obtained via the proposed model was evaluated using a ranging algorithm and classified using a Support Vector Machine (SVM) and Naive Bayes (NB) classifier. More recent research utilized a deep neural network (DNN) based model to learn the latent semantic features to improve the generalization capabilities and overall accuracy in NED. A DNN-based approach was implemented to extract the entities to construct Health-KGs, which used the inverse document frequency-based filtering [144] while a similar idea was proposed in another research, but a Co-training algorithm was used. The idea was to train two different classifiers independently and use independent target instance distribution [162].

5.5.3. Entity Linkage

Entity Linkage (NEL), also known as entity resolution or record linkage, is a process in Health-KGs that involves identifying and linking related information about entities, such as patients, medical conditions, or treatments, across different data sources or within the same dataset [163,164]. The goal is to connect information that refers to the same real-world entity, even if the data is represented differently or contains variations [165].

Numerous works have been published to extract clinical concepts from reports. In this work, we particularly focus on the techniques and models that focus on extracting the Unified Medical Language System (UMLS) through NER and NEL. The state-of-the-art models like GloVe [166], BERT [39], ELMo [146], word2vec [167], and fastText [168] are popularly used to identify UMLS concepts. In most published literature, the BERT-based contextualized embedding approach has outperformed other techniques. The other techniques BlueBERT [169], BioBERT [115], SapBERT [170], UmlsBERT [171], NeighBERT [172] have utilized UMLS and their synonyms for EL. Although there has been significant success in the past, most of the published work suffers from numerous entity-linking errors attributed to unstructured text snippets and the presence of multiple relationships associated with entities [173]. Pretraining corpus, not architecture, is the dominant driver of clinical performance: models exposed to in-domain clinical notes (ClinicalBERT, BlueBERT) lead extraction tasks, while UMLS-aligned models (SapBERT, UmlsBERT, NeighBERT) lead entity linking. Table 6 summarizes these clinical and biomedical language models by pretraining corpus and primary use in Health-KGs.

Table 5. Representative LLM-based systems for Health-KG construction and reasoning. Grouped by primary role of the knowledge graph. Rows marked “(general)” are domain-agnostic paradigms adapted to the clinical setting.

System [Ref]	Year	Primary task	Role of KG & grounding mechanism	Data / benchmark	Reported result / finding
Agrawal et al. [174]	2022	Extraction	Schema-in-prompt zero/few-shot clinical information extraction	Clinical notes (i2b2/CASI)	Few-shot LLM IE competitive without task-specific training
SAC-KG [175]	2024	Construction	LLM Generator + rule-based Verifier + fine-tuned Pruner	Domain corpora	Automated domain-KG construction with staged error control
KARMA [176]	2025	Construction / enrichment	Nine collaborative agents; schema alignment + conflict resolution	1,200 PubMed articles	~38k new entities; 83.1% LLM-verified correctness
Yang et al. [177]	2025	Construction (EHR)	GPT-4 entity recognition into Neo4j; de-identification pre-step	Multicenter sepsis EHR (~10.5k patients)	Prompt-engineered multicenter clinical KG for decision support
Think-on-Graph [178]	2023	Reasoning (general)	Interleaved LLM reasoning with explicit KG traversal	General KGs	Interpretable multi-hop inference chains
GraphRAG [179]	2024	Reasoning / QA (general)	Corpus indexed as KG; community-level summaries	General corpora	Global, sense-making queries beyond context window
Gao et al. [180]	2023	Diagnosis prediction	UMLS 2-hop subgraph injected into LLM	MIMIC-III	KG grounding improves diagnosis prediction
RAM-EHR [181]	2024	Clinical prediction	Retrieval augmentation over multi-source EHR	EHR predictive tasks	Retrieval augmentation improves predictive accuracy
MedRAG [182]	2025	Diagnosis / reasoning	KG-elicited reasoning over a four-tier diagnostic KG	DDXPlus + private CPDD	Lower misdiagnosis; proactive follow-up questions
MedGraphRAG [183]	2025	QA / reasoning	Triple Graph Construction + U-Retrieval; evidence linking	Medical QA benchmarks	Evidence-based, source-grounded responses
AMG-RAG [184]	2025	Question answering	Agentic auto-construction/updating of MKG + external retrieval	MedQA, MedMCQA	MedQA F1 ≈74.1% with an 8B backbone
HyKGE [185]	2024	QA / reasoning	Hypothesis-KG-enhanced retrieval	Medical QA	Improved accuracy and answer reliability
TC-RAG [186]	2024	Reasoning	Turing-complete/stateful RAG with a memory stack	Medical LLM systems	Controlled retrieval state for deeper multi-hop
PrimeKG [187]	2023	Resource (KG)	Openly available precision-medicine KG substrate	Drugs/diseases/phenotypes/genes	Reusable multimodal biomedical KG for grounding

6. Large Language Models for Health-KG Construction and Reasoning

The construction pipeline in Section IV has historically depended on task-specific extractors, translational or neural embedding models, and rule- or machine-learning-based linkers, each requiring substantial annotated data and domain engineering. The emergence of instruction-tuned large language models (LLMs) has reshaped every stage of this pipeline. In contrast to the encoder-only models discussed earlier, such as BERT [39] and its clinical variants, generative LLMs perform extraction, normalization, and multi-hop reasoning with few or no task-specific labels. The resulting relationship characterized as bidirectional: Health-KGs supply LLMs with verifiable, structured clinical knowledge and provenance, while LLMs lower the cost of constructing, completing, and querying Health-KGs [50]. This section organizes recent work along three threads: LLM-driven extraction for construction (VI-A), retrieval- and graph-augmented clinical reasoning that couples Health-KGs with LLMs (VI-B), and the use of LLMs as clinical knowledge sources together with the factuality challenge that constrains their autonomous use (VI-C). Figure 4 presents the end-to-end Health-KG construction-and-application pipeline, in which LLM- and RAG-based components augment extraction, linking, and reasoning while completed knowledge feeds back for continual update. Table 5 summarizes representative LLM-based systems for Health-KG construction and reasoning, organized by the role the graph plays and the grounding mechanism employed.

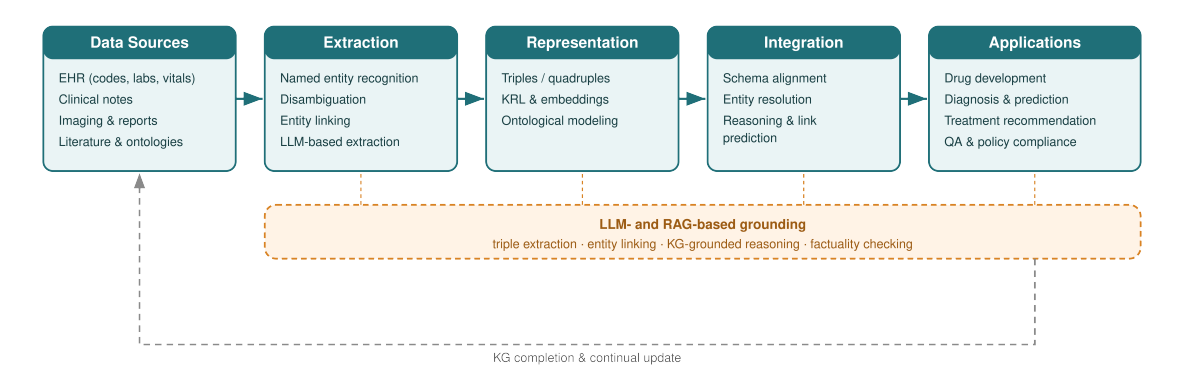


Figure 4. End-to-end Health-KG construction and application pipeline. LLM- and RAG-based components (lower band) now augment extraction, linking, and reasoning, while completed knowledge feeds back for continual update.

6.1. LLM-Driven Entity and Relation Extraction

The named-entity recognition, disambiguation, and linking methods reviewed in Section V frame extraction as a sequence of supervised classification tasks. Generative LLMs reformulate extraction as a conditional generation problem, in which entities, relations, and their normalized identifiers are produced directly from clinical text under natural-language or schema-constrained prompts. This reformulation substantially reduces the annotation burden that earlier supervised pipelines required, and in-context exemplars allow rapid adaptation across note types and institutions without retraining. The output can be constrained to controlled vocabularies, mapping extracted mentions to UMLS [22], SNOMED-CT [23], or ICD codes [24] during decoding, which integrates the recognition and linking stages that were previously separate. These gains come with characteristic failure modes: LLMs may hallucinate non-existent entities, assign inconsistent semantic types, or produce identifiers absent from the target ontology. Consequently, LLM-based extraction for Health-KG construction is most reliable when paired with ontology grounding and human-in-the-loop validation, trading the label efficiency of generative extraction against the precision and controllability of the supervised neural models.

Table 6. Clinical/biomedical language models for Health-KG extraction and linking.

Model	Pretraining Corpus	Primary Use in Health-KGs
SciBERT [112]	Scientific (Semantic Scholar)	Scientific-text NER / classification
BioBERT [115]	PubMed abstracts + PMC full text	Biomedical NER, RE, QA
ClinicalBERT [114]	MIMIC-III clinical notes (over BERT/BioBERT)	Clinical NER, readmission
BlueBERT [115]	PubMed + MIMIC-III	Biomedical + clinical benchmarks (BLUE)
SapBERT [170]	UMLS synonym pairs (self-alignment)	Entity linking / concept normalization
UmlsBERT [171]	+ UMLS Metathesaurus relations	Clinical NER, NLI
NeighBERT [172]	Relation-induced dense retrieval	Medical entity linking

6.2. Retrieval and Graph-Augmented Clinical Reasoning

The parametric LLMs cannot reliably access or update factual knowledge, a complementary line of work grounds their outputs in external structured knowledge. Retrieval-augmented generation (RAG) [188] retrieves relevant passages from EHR corpora, clinical guidelines, or literature and conditions generation on them, providing provenance and allowing knowledge to be updated without retraining. Graph-structured variants extend this idea to Health-KGs. GraphRAG [179] indexes a corpus as a knowledge graph and uses community-level summaries to answer global, sense-making queries that

exceed an LLM’s context window, a capability well suited to longitudinal, multi-document patient records. Think-on-Graph [178] interleaves LLM reasoning with explicit traversal over a knowledge graph, yielding interpretable multi-hop inference chains rather than opaque parametric recall. These paradigms have been instantiated in clinical settings: retrieval augmentation has been applied directly to predictive tasks over electronic health records [181], and a medical knowledge graph has been injected into an LLM to improve diagnosis prediction on MIMIC-III [180]. The consistent finding across these systems is that coupling an LLM with a Health-KG improves factuality and interpretability and supplies the long-tail and temporally evolving knowledge that parametric models omit. The open problems are retrieval quality, schema alignment between the LLM’s outputs and the target KG, and the stability of multi-hop traversal as reasoning depth increases.

6.3. LLMs as Clinical Knowledge Sources and the Factuality Challenge

The final thread treats the LLM itself as a partial knowledge source [189] showed that large models encode considerable clinical knowledge and that comprehension and reasoning scale with model size and instruction tuning, while also demonstrating that performance remains below that of clinicians and that automated benchmarks understate safety-critical errors. This gap defines the principal barrier to using LLMs as autonomous constructors or curators of Health-KGs: hallucination, in which fluent but unsupported facts are generated, is difficult to detect with standard benchmarks [190], and an unsupported triple introduced into a clinical KG can propagate into downstream decision support. Knowledge-graph grounding, retrieval with provenance, and explicit verification are therefore not optional refinements but prerequisites for safe deployment, a point reinforced by surveys of RAG mitigation strategies [191]. These factuality requirements also intersect with the compliance [192,193] and privacy concerns of Section VII-D, since clinical text placed in LLM prompts constitutes protected health information, motivating on-premises, federated, or de-identified construction workflows. In aggregate, LLMs shift the bottleneck of Health-KG construction from extraction labor toward verification, grounding, and governance, and the absence of standardized clinical KG–LLM benchmarks and factuality guarantees remains the central open challenge that Section VIII develops further.

7. Application of Health-KGs

Health-KGs, an EHR-curated knowledge graph, is a powerful tool for representing complex real-world medical knowledge to assist various clinical decision support applications. In this contribution, several Health-KGs applications, i.e., drug development, clinical trial, diagnosis and predictions, treatment recommendation, question answering, and health policy compliance, will be covered. Table 7 summarizes representative Health-KG applications together with their methods, datasets, and evaluation metrics.

Table 7. Health-KGs Application; Ref: References, YOP: Year of Publication; EM: Evaluation Metrics; AUC: Area Under Curve, AUPR: Area Under the Precision-Recall curve; MCC: Matthews Correlation Coefficient; MAE: Mean Absolute Error; ACC: Accuracy; N.A: Not Applicable.

Ref	Applications	Methods	Datasets	EM
[194]	Discover candidate drugs	SemaTyP	Metamap, PubMed, Sem-Rep, UMLS	Precision, Recall, F-score
[116]	Discover candidate drugs	GrEDeL	TT, PubMed UMLS	Precision, Recall, F-score
[195]	Drug-drug interactions	ML, C-LSTM	KEGG, DrugBank, & PharmGKB	AUPR, MCC, AUC, F-score, MAE
[118]	Drug-drug interactions	LR, NB, RF	PharmGKB, DrugBank v4 & v5, KEGG	AUPR, F-score, AUC
[196]	Diagnosis prediction	KAME	MIMIC-III	Accuracy, Precision
[197]	Diagnosis prediction	GNDP	MIMIC-III, MIMIC-IV	Accuracy, Precision

Table 7. Cont.

Ref	Applications	Methods	Datasets	EM
[198]	Diagnosis prediction	GRU	MIMIC-III, MIMIC-IV	Accuracy, Precision
[180]	Diagnosis prediction	LLM	MIMIC-III	Recall, Precision, F-Score
[199]	Diagnosis prediction & disease progression	RLM	MIMIC-III, MIMIC-IV	AUC
[200]	Treatment recommendation	JNN	Shanghai Hospital (Private)	Recall, AP
[201]	Question answering	MKHAN	Dingxiang & Chunyu Doctor	Precision
[202]	Drug repurposing	KG-Predict	MGI, GOA, GTEx	N.A
[203]	Drug recommendation	KGCNN+ SA	Kaggle Open-source	AUC, MAE, MSE, RMS
[204]	Health policy compliance	TCEF	Synthetic	Elapsed Time
[46]	Health policy compliance	KGF	Private	N.A.

7.1. Drug Development & Clinical Trials

Health KGs have been popularly implemented in drug development and drug repurposing. In a recent ML-based approach [195], integrated KGs were used to predict drug-drug interactions utilizing the features from KEGG [205], DrugBank [25], and PharmGKB [26]. The default parameters provided in the open-source implementations of PBG9, SimpleIE12, TransE11, and CrossE10 were utilized for knowledge graph embedding. The complex embedding method based on the Convolutional-LSTM (C-LSTM) network was reported as the optimal performance model. In another work involving a similar objective, an end-to-end framework, KG-neural network (KGNN), was proposed to capture high-order structures and semantic relations [206]. The other work explored the discovery of candidate drugs for diseases of interest by creating Health-KGs from the therapeutic target database (TT), PubMed database, and UMLS. The authors proposed the LSTM-based model GrEDeL and compared it using Random Forest and Support Vector Machine [116], while in another literature to resolve similar objective, a logistic regression model called SemaTyP was trained to understand the semantic types of paths associated with known drug therapies within the Health-KGs [194]. The acquired model is applied to unveil drug therapies for novel diseases. SemaTyP model was compared with the two Random walk algorithm (RWA)-based methods [207,208]. In another study to predict DDI, authors utilized Logistic Regression (LR), Naive Bayes (NB), and Random Forest (RF) in conjunction with TransD, TransE, and RDF2Vec on the DrugBank knowledge graph, employing a 10-fold traditional cross-validation approach. The model was evaluated RDF2Vec on diverse drug knowledge graphs, including DrugBank, PharmGKB, and KEGG, to predict previously unknown drug-drug interactions [118].

To enhance drug discovery efficiency, scientists investigate the repurposing of ‘existing’ drugs [202]. In a recent study, KG-Predict was proposed and constructed using a stack of CompGCN layers. The Mouse Genome Informatics (MGI) database [209], Gene Ontology Annotation (GOA) database [210], and Genotype-Tissue Expression (GTEx) [211] were used. This approach involves mitigating risks associated with advancing to clinical phases earlier. Evaluating Adverse Drug Reactions (ADRs) in pharmacological analysis and clinical application is crucial. Consequently, during the stages of safety pharmacology analysis and phase III clinical trials, researchers often allocate substantial resources to assess ADRs by identifying side effects and Drug-Drug Interactions (DDIs).

7.2. Diagnosis Prediction & Treatment Recommendation

Health-KGs with EHR data from different sources have been widely used in health care recommendation systems for diagnosis prediction & treatment or drug recommendation. Diagnosis Prediction early is a significant research area in medical informatics because it can anticipate patient prognosis and mitigate preventable adverse events. In a recently published study [197,199], Health-

KGs was built using the PLAGH dataset, the MIMIC-III dataset [20], and the MIMIC-IV dataset [21]. The Health-KGs were derived from patients' current diseases or risk factors and a disease entity representing the predicted ailment. The diagnosis prediction and disease progression were achieved via a policy-based reinforcement learning module (RLM) [199]. Besides, the other study constructed sequential patient graphs and utilized the Graph Neural Network-Based Diagnosis Prediction (GNDP) for diagnosis prediction [197]. In a similar goal of diagnosis prediction task, knowledge of the medical graph (KAME) [196] an end-to-end, accurate, and robust recurrent neural networks (RNNs) with attention mechanism was implemented and large Language Models (LLM) for Diagnosis Prediction [180] and gated recurrent unit (GRU) was proposed [198]. The performance of the KAME & LLM model in comparison to the existing state-of-the-art improved significantly [196].

A recent study used a graph-based KG-Convolutional Neural Network (KGCNN) coupled with sentiment analysis (SA) extracted from the public reviews on drugs to develop a drug recommendation system. The Health-KGs developed captured user-item relatedness by mining its associated attributes effectively [203]. In another work, a thyroid treatment recommendation framework was proposed, which utilized TransD embedding and provided the treatment recommendation using a joint neural model (JNN). The output is ranked using a classifier trained on internal EMR data collected from Shanghai Hospital [200].

7.3. Question Answering

Question answering within the medical domain is crucial as a decision support tool. An EMR-based knowledge graph specific to the domain ensures the provision of intelligent and contextually appropriate answers to patient-specific queries, surpassing the capabilities of a conventional schema-based question-answer system [77]. The significance of comprehending the semantics and context of data retrieval for information retrieval, especially when seeking domain-specific information, cannot be overstated. In recent studies, knowledge graph-based question answering (KGQA) was proposed to extract named entities and relations from data. The related entities obtained were ranked using weighted path ranking on the knowledge graph [212]. In another study, the Multi-Modal Knowledge-aware Hierarchical Attention Network (MKHAN) was developed to efficiently harness a multi-modal knowledge graph (MKG) for transparent medical question answering. MKHAN utilized the underlying sequential dependencies within a path extracted from MKG to deduce the underlying rationale of question-answer interactions [201].

7.4. Health Policy Compliance

The significance of data in scientific experiments within healthcare is invaluable, particularly given the exponential growth in data generation. In the healthcare sector, the Health Insurance Portability and Accountability Act (HIPAA) is a pivotal privacy regulation dedicated to protecting patient data in the United States [193,213]. Comparable privacy regulations exist in other domains, exemplified by the General Data Protection Regulation (GDPR), which outlines companies' restrictions on using customer information [192,214]. To ensure data privacy, data providers institute security protocols that precisely outline criteria for obtaining, analyzing, and sharing personal and confidential information during the integration process. A recent work proposed an integrated KG-based framework for protecting privacy data across healthcare providers. The framework prevented excess or unwanted access using Attribute Based Access Control (ABAC) [204]. In another study, the Trusted Compliance Enforcement Framework (TCEF) was developed using semantic web technologies where a trust Score is assigned to each participant involved in the data exchange process, incorporating ontologies alongside policy reasoners to guarantee that data access aligns with health regulations such as HIPAA [215,216]. In modern wearable fitness devices, data such as heart rate, glucose and stress levels, medication adherence, menstrual cycle, and body metabolic ratio are measured. Health data captured outside of hospital settings often operates beyond compliance regulations. A recent study suggested a KG-based framework (KGF) to scrutinize the privacy policies governing the data and information utilized by wearable devices [46].

8. Evaluation, Open Challenges and Future Direction

8.1. Evaluation and Data-Centric Challenges

Health-KG methods are evaluated with task-specific and largely incompatible protocols: link prediction reports MRR and Hits@k; extraction reports precision, recall, and F1; and downstream clinical tasks report AUROC and AUPRC. Because each study constructs its own graph from a different slice of MIMIC or a private cohort, results are rarely comparable across papers, and improvements cannot be attributed cleanly to representation, extraction, or integration. The absence of a standardized, openly released Health-KG benchmark with fixed splits, a shared schema, and harmonized evaluation tasks is itself a finding of this review and a prerequisite for measurable progress. A community benchmark built on credentialed but openly governed resources (e.g., a MIMIC-IV-derived graph with canonical link-prediction, NER, entity-linking, and diagnosis-prediction tracks) would let the field compare embeddings, extractors, and LLM-grounded systems on equal footing.

Beyond evaluation, the underlying data pose challenges that any such benchmark must confront. EHR data are heterogeneous, noisy, sparse, and irregularly sampled, and they mix structured codes with free text and images, which complicates the clean triple/quadruple abstraction on which most Health-KGs rely. Missingness is informative (a missing lab is not missing at random), and the same clinical concept appears under multiple codes and surface forms; addressing this requires imputation-aware knowledge-representation learning, multimodal fusion at the representation stage [65,66], and quadruple schemas that attach provenance and confidence to each fact so that downstream reasoning can discount low-confidence edges. Cross-institutional integration compounds the problem, since divergent terminologies, local code systems, and schema mismatch mean that equivalent concepts are encoded differently across sites and naive merging produces conflicting or duplicate facts. Ontology matching and schema alignment [128,133,134], combined with LLM-assisted concept normalization validated against UMLS/SNOMED-CT and multiview integration methods that tolerate partial overlap, offer a path forward, as demonstrated for cross-institutional COVID-19 harmonization [137].

Two further data-centric axes remain open. Temporally, clinical knowledge and patient state evolve, yet most Health-KGs are static snapshots: relations are valid only over intervals (a medication is current, then discontinued), and disease trajectories depend on order and timing, motivating temporal knowledge-graph completion [78], time-stamped quadruples, and continual/streaming construction that updates the graph as new encounters arrive rather than rebuilding from scratch. In terms of modality, most Health-KGs are text- and code-centric, omitting imaging and omics; aligning continuous signals (radiology, pathology, genomics) with discrete graph entities requires shared representation spaces and multimodal encoders that project image- and omics-derived features into the same embedding space as clinical concepts [67], enabling a single graph to support image-grounded reasoning.

8.2. Trustworthy, Privacy-Preserving, and Reproducible Construction

EHR-derived graphs contain protected health information, and LLM-based construction risks exposing it through prompts. HIPAA and GDPR [213,214] constrain data movement, while useful KGs require cross-site information; reconciling these pressures calls for federated knowledge-graph embedding, on-premises or de-identified LLM extraction, synthetic-data augmentation, and policy-aware access control such as the compliance frameworks discussed in Section VII-D [46,204,215].

Trust concerns extend from privacy to factuality as LLMs enter the construction loop, where factuality and reproducibility become first-order concerns. Hallucinated facts can silently corrupt a clinical KG [190], and few studies release code, graphs, or prompts, leaving reporting inconsistent. Mandatory provenance on every asserted fact, automated verification against trusted ontologies, and adoption of reporting standards are therefore needed so that construction and evaluation are reproducible.

9. Conclusions

This review examined the construction and application of Health Knowledge Graphs from electronic health records along three axes—representation, extraction, and integration—which keeping the review organized the field through a unified taxonomy and an explicit positioning against prior reviews. We reviewed knowledge-representation learning and graph-embedding models, entity recognition, disambiguation, and linking, and the integration methods that reconcile heterogeneous clinical sources, together with the principal EHR datasets and clinical language models that underpin them. We then synthesized the emerging role of large language models in Health-KG construction and KG-grounded clinical reasoning, including retrieval- and graph-augmented generation and the factuality constraints that govern their safe use. Across drug development, diagnosis and treatment recommendation, question answering, and policy compliance, Health-KGs consistently improve the structure, provenance, and interpretability of clinical decision support. We argued that the central obstacles to measurable progress are the absence of a standardized, openly governed Health-KG benchmark and the intertwined demands of temporal modeling, multimodal fusion, privacy-preserving construction, and factuality verification, which we identify as priorities for future work.

Acknowledgments: This work was partially funded by the NSF Award Number DMS2436549, Collaborative Research: FDT-BioTech: Aspects of Digital Twin Studies for Neuroimages.

References

1. Nicholson, D.N.; Greene, C.S. Constructing knowledge graphs and their biomedical applications. *Computational and Structural Biotechnology Journal* **2020**, *18*, 1414–1428.
2. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. *IEEE Transactions on Neural Networks and Learning Systems* **2020**, *33*, 494–514.
3. Abu-Salih, B. Domain-specific knowledge graphs: A survey. *of Network and Computer Applications* **2021**, *185*.
4. Mrdjenovich, D.; Horton, M.K.; Montoya, J.H.; Legaspi, C.M.; Dwaraknath, S.S.; Tshitoyan, V.; Jain, A.; Persson, K.A. propnet: A Knowledge Graph for Materials Science. *Matter* **2020**.
5. Dong, J.; Jing, X.; Lu, X.; Liu, J.; Li, H.; Cao, X.; Du, C.; Li, J.; Li, L. Process knowledge graph modeling techniques and application methods for ship heterogeneous models. *Scientific Reports* **2021**, *12*.
6. Huang, Z.; Yang, J.; Harmelen, F.V.; Hu, Q. Constructing Knowledge Graphs of Depression. In Proceedings of the International Conference on Health Information Science, 2017.
7. Zeng, X.; Tu, X.; Liu, Y.; Fu, X.; Su, Y. Toward better drug discovery with knowledge graph. *Current opinion in structural biology* **2021**, *72*, 114–126.
8. Xiong, Z.; Huang, F.; Wang, Z.; Liu, S.; Zhang, W. A Multimodal Framework for Improving in Silico Drug Repositioning With the Prior Knowledge From Knowledge Graphs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **2021**, *19*, 2623–2631.
9. Rotmensch, M.; Halpern, Y.; Tlimat, A.; Horng, S.; Sontag, D.A. Learning a Health Knowledge Graph from Electronic Medical Records. *Scientific Reports* **2017**, *7*.
10. Cai, Q.; Wang, H.; Li, Z.; Liu, X. A Survey on Multimodal Data-Driven Smart Healthcare Systems: Approaches and Applications. *IEEE Access* **2019**, *7*, 133583–133599.
11. Wang, P.; Jiang, H.; Xu, J.; Zhang, Q. Knowledge Graph Construction and Applications for Web Search and Beyond. *Data Intelligence* **2019**, *1*, 333–349.
12. Wang, X.; He, X.; Cao, Y.; Liu, M.; Chua, T.S. KGAT: Knowledge Graph Attention Network for Recommendation. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* **2019**.
13. Lin, B.Y.; Chen, X.; Chen, J.; Ren, X. KagNet: Knowledge-Aware Graph Networks for Commonsense Reasoning. *ArXiv* **2019**, *abs/1909.02151*.
14. Chen, X.; Jia, S.; Xiang, Y. A review: Knowledge reasoning over knowledge graph. *Expert Syst. Appl.* **2020**, *141*.
15. Kan, X.; Cui, H.; Yang, C. Zero-Shot Scene Graph Relation Prediction through Commonsense Knowledge Integration. In Proceedings of the ECML/PKDD, 2021.
16. Kim, E.; Rubinstein, S.M.; Nead, K.T.; Wojcieszynski, A.P.; Gabriel, P.E.; Warner, J.L. The Evolving Use of Electronic Health Records (EHR) for Research. *Seminars in radiation oncology* **2019**, *29* 4, 354–361.

17. Alammari, D.; Banta, J.E.; Shah, H.; Reibling, E.T.; Ramadan, M.M. Meaningful Use of Electronic Health Records and Ambulatory Healthcare Quality Measures. *Cureus* **2021**, *13*.
18. Bouayad, L.; Ialynytchev, A.; Padmanabhan, B. Patient Health Record Systems Scope and Functionalities: Literature Review and Future Directions. *Journal of Medical Internet Research* **2017**, *19*.
19. Rajkomar, A.; Oren, E.; Chen, K.; Dai, A.M.; Hajaj, N.; Hardt, M.; Liu, P.J.; Liu, X.; Marcus, J.; Sun, M.; et al. Scalable and accurate deep learning with electronic health records. *NPJ Digital Medicine* **2018**, *1*.
20. Johnson, A.E.W.; Pollard, T.J.; Shen, L.; wei H. Lehman, L.; Feng, M.; Ghassemi, M.M.; Moody, B.; Szolovits, P.; Celi, L.A.; Mark, R.G. MIMIC-III, a freely accessible critical care database. *Scientific Data* **2016**, *3*.
21. Johnson, A.E.W.; Bulgarelli, L.; Shen, L.; Gayles, A.; Shammout, A.; Horng, S.; Pollard, T.J.; Moody, B.; Gow, B.; wei H. Lehman, L.; et al. MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data* **2023**, *10*.
22. Bodenreider, O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic acids research* **2004**, *32 Database issue*, D267–70.
23. Donnelly, K.P. SNOMED-CT: The advanced terminology and coding system for eHealth. *Studies in health technology and informatics* **2006**, *121*, 279–90.
24. Quan, H.; Sundararajan, V.; Halfon, P.; andrew. fong.; Burnand, B.; Luthi, J.C.; Saunders, L.D.; Beck, C.A.; Feasby, T.E.; Ghali, W.A. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. *Medical Care* **2005**, *43*, 1130–1139.
25. Wishart, D.S.; Knox, C.; Guo, A.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Research* **2007**, *36*, D901 – D906.
26. Thorn, C.F.; Klein, T.E.; Altman, R.B. PharmGKB: the Pharmacogenomics Knowledge Base. *Methods in molecular biology* **2013**, *1015*, 311–20.
27. Kanehisa, M. The KEGG database. *Novartis Foundation symposium* **2002**, *247*, 91–101; discussion 101–3, 119–28, 244–52.
28. Pratiher, S.; Chatteraj, S. Diving Deep onto Discriminative Ensemble of Histological Hashing & Class-Specific Manifold Learning for Multi-class Breast Carcinoma Taxonomy. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* **2018**, pp. 1025–1029.
29. Pratiher, S.; Chatteraj, S.; Nawn, D.; Pal, M.; Paul, R.R.; Konik, H.; Chatterjee, J. A Multi-Scale Context Aggregation Enriched MLP-Mixer Model for Oral Cancer Screening from Oral Sub-epithelial Connective Tissues. *2022 30th European Signal Processing Conference (EUSIPCO)* **2022**, pp. 1323–1327.
30. Pratiher, S.; Chatteraj, S.; Agarwal, S.K.; Bhattacharya, S. Grading Tumor Malignancy via Deep Bidirectional LSTM on Graph Manifold Encoded Histopathological Image. *2018 IEEE International Conference on Data Mining Workshops (ICDMW)* **2018**, pp. 674–681.
31. Kumar, S.; Agarwal, A.; Golla, S.; Tanamala, S.; Upadhyay, U.; Chatteraj, S.; Putha, P.; Chilamkurthy, S. Mind the Clot: Automated LVO Detection on CTA using Deep Learning. *2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)* **2023**, pp. 2495–2504.
32. Govindarajan, A.; Govindarajan, A.; Tanamala, S.; Chatteraj, S.; Reddy, B.; Agrawal, R.; Iyer, D.; Srivastava, A.; Kumar, P.; Putha, P. Role of an Automated Deep Learning Algorithm for Reliable Screening of Abnormality in Chest Radiographs: A Prospective Multicenter Quality Improvement Study. *Diagnostics* **2022**, *12*.
33. Feng, J.; Lee, J.; Vesoulis, Z.; Li, F. Predicting mortality risk for preterm infants using deep learning models with time-series vital sign data. *NPJ Digital Medicine* **2021**, *4*.
34. Nawn, D.; Pratiher, S.; Chatteraj, S.; Chakraborty, D.; Pal, M.; Paul, R.R.; Dutta, S.; Chatterjee, J. Multifractal Alterations in Oral Sub-Epithelial Connective Tissue During Progression of Pre-Cancer and Cancer. *IEEE Journal of Biomedical and Health Informatics* **2020**, *25*, 152–162.
35. Blake, S.R.; Das, N.; Tadepalli, M.; Reddy, B.; Singh, A.; Agrawal, R.; Chatteraj, S.; Shah, D.; Putha, P. Using Artificial Intelligence to Stratify Normal versus Abnormal Chest X-rays: External Validation of a Deep Learning Algorithm at East Kent Hospitals University NHS Foundation Trust. *Diagnostics* **2023**, *13*.
36. Blom, M.C.; Ashfaq, A.; Sant'Anna, A.; Anderson, P.D.; Lingman, M. Training machine learning models to predict 30-day mortality in patients discharged from the emergency department: a retrospective, population-based registry study. *BMJ Open* **2019**, *9*.
37. Courtright, K.R.; Chivers, C.; Becker, M.; Regli, S.H.; Pepper, L.; Draugelis, M.; O'Connor, N.R. Electronic Health Record Mortality Prediction Model for Targeted Palliative Care Among Hospitalized Medical Patients: a Pilot Quasi-experimental Study. *Journal of General Internal Medicine* **2019**, *34*, 1841 – 1847.

38. Levin, S.R.; Barnes, S.L.; Toerper, M.F.; Debraine, A.; DeAngelo, A.; Hamrock, E.; Hinson, J.S.; Hoyer, E.H.; Dungarani, T.; Howell, E.E. Machine-learning-based hospital discharge predictions can support multidisciplinary rounds and decrease hospital length-of-stay. *BMJ Innovations* **2020**, *7*, 414 – 421.
39. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the North American Chapter of the Association for Computational Linguistics, 2019.
40. Brown, T.B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language Models are Few-Shot Learners. *ArXiv* **2020**, *abs/2005.14165*.
41. Yan, J.; Raman, M.; Chan, A.; Zhang, T.; Rossi, R.A.; Zhao, H.; Kim, S.; Lipka, N.; Ren, X. Learning Contextualized Knowledge Structures for Commonsense Reasoning. In Proceedings of the Findings, 2020.
42. Yu, H.; Li, K.; Dong, W.; Song, S.; Gao, C.; Shi, J. Attention-based cross domain graph neural network for prediction of drug-drug interactions. *Briefings in bioinformatics* **2023**.
43. Wang, S.; Du, Z.; Ding, M.; Rodriguez-Paton, A.; Song, T. KG-DTI: a knowledge graph based deep learning method for drug-target interaction predictions and Alzheimer's disease drug repositions. *Applied Intelligence* **2021**, *52*, 846–857.
44. Mehta, N.; Pandit, A. Concurrence of big data analytics and healthcare: A systematic review. *International journal of medical informatics* **2018**, *114*, 57–65.
45. Dash, S.; Shakyawar, S.K.; Sharma, M.; Kaushik, S.R. Big data in healthcare: management, analysis and future prospects. *Journal of Big Data* **2019**, *6*.
46. Echenim, K.U.; Elluri, L.; Pande, K. Ensuring Privacy Policy Compliance of Wearables with IoT Regulations. 2023.
47. Chatteraj, S.; Reddy, B.; Tadepalli, M.; Putha, P. Comparing deep learning models for tuberculosis detection: A retrospective study of digital vs. analog chest radiographs. *The Indian journal of tuberculosis* **2024**, *72 Suppl 2*, S43–S46.
48. Page, M.J.; Moher, D.; Bossuyt, P.M.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *The BMJ* **2020**, 372.
49. Silva, M.C.; Eugénio, P.; Faria, D.; Pesquita, C. Ontologies and Knowledge Graphs in Oncology Research. *Cancers* **2022**, *14*.
50. Pan, S.; Luo, L.; Wang, Y.; Chen, C.; Wang, J.; Wu, X. Unifying Large Language Models and Knowledge Graphs: A Roadmap. *IEEE Transactions on Knowledge and Data Engineering* **2023**, *36*, 3580–3599.
51. Hirsch, J.A.; Leslie-Mazwi, T.M.; Nicola, G.N.; Barr, R.; Bello, J.A.; Donovan, W.; Tu, R.; Alson, M.D.; Manchikanti, L. Current procedural terminology; a primer. *Journal of NeuroInterventional Surgery* **2014**, *7*, 309 – 312.
52. Gr, B. International statistical classification of diseases and related health problems. Tenth revision. *World health statistics quarterly. Rapport trimestriel de statistiques sanitaires mondiales* **1988**, *41 1*, 32–6.
53. Bodenreider, O. Biomedical Ontologies in Action: Role in Knowledge Management, Data Integration and Decision Support. *Yearbook of Medical Informatics* **2008**, *17*, 67 – 79.
54. Henry, S.; Buchan, K.; Filannino, M.; Stubbs, A.; Uzuner, Ö. 2018 N2c2 Shared Task on Adverse Drug Events and Medication Extraction in Electronic Health Records. *Journal of the American Medical Informatics Association : JAMIA* **2020**.
55. Denecke, K. Biomedical Standards and Open Health Data. 2020.
56. Wulff, A.; Haarbrandt, B.; Tute, E.; Marschollek, M.; Beerbaum, P.; Jack, T. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artificial intelligence in medicine* **2018**, *89*, 10–23.
57. Sedano, F.J.F.; Cuadrado, M.T.; Clemente, Y.C.; Serrano-Balazote, P.; Cano, D.M.; Viejo, M.R. Patient Summary and medicines reconciliation: application of the ISO/CEN EN 13606 standard in clinical practice. *Studies in health technology and informatics* **2011**, *166*, 189–96.
58. Smith, B.; Ceusters, W. HL7 RIM: An Incoherent Standard. *Studies in health technology and informatics* **2006**, *124*, 133–8.
59. Vorisek, C.N.; Lehne, M.; Klopfenstein, S.A.I.; Mayer, P.J.; Bartschke, A.; Haese, T.; Thun, S. Fast Healthcare Interoperability Resources (FHIR) for Interoperability in Health Research: Systematic Review. *JMIR Medical Informatics* **2021**, *10*.
60. Fernández-breis, J.T.; Vicente, P.J.V.; Tortosa, M.M.; Moner, D.; Maldonado, J.A.; Valencia-García, R.; Miranda-Mena, T.G. Using semantic technologies to promote interoperability between electronic healthcare records'

- information models. *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* **2006**, pp. 2614–2617.
61. Lewalle, P.; Rodrigues, J.M.; Zanstra, P.E.; Üstün, B.; Kalra, D.; Surján, G.; Rector, A.L.; Stroetmann, V.N.; Virtanen, M. A Deployment and Research Roadmap for Semantic Interoperability: the EU SemanticHEALTH project. *Studies in health technology and informatics* **2008**, *136*, 635–40.
 62. Pollard, T.J.; Johnson, A.E.W.; Raffa, J.D.; Celi, L.A.; Mark, R.G.; Badawi, O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Scientific Data* **2018**, *5*.
 63. Li, P.; Xie, C.; Pollard, T.; Lodano, C.; Celi, L.A.; He, K.; et al. Promoting secondary analysis of electronic medical records in China: summary of the PLAGH-MIT critical data conference and health datathon. *JMIR Medical Informatics* **2017**, *5*, e43.
 64. Zhao, H.; Cao, X.; Zhang, J.; Wang, Y.; He, K. Construction of a non-mutually exclusive decision tree for medication recommendation of chronic heart failure. *Frontiers in Pharmacology* **2022**, *12*, 758573.
 65. Zhao, X.; Jia, Y.; Li, A.; Jiang, R.; Xie, H.; Song, Y.; Han, W. Multi-source knowledge fusion: a survey. *World Wide Web* **2019**, *23*, 2567–2592.
 66. Lin, Y.; Lu, K.; Yu, S.; Cai, T.; Zitnik, M. Multimodal Learning on Graphs for Disease Relation Extraction. *Journal of biomedical informatics* **2022**, p. 104415.
 67. Zhu, C.; Yang, Z.H.; Xia, X.; Li, N.; Zhong, F.; Liu, L. Multimodal reasoning based on knowledge graph embedding for specific diseases. *Bioinformatics* **2022**, *38*, 2235 – 2245.
 68. Tsai, Y.H.H.; Bai, S.; Liang, P.P.; Kolter, J.Z.; Morency, L.P.; Salakhutdinov, R. Multimodal Transformer for Unaligned Multimodal Language Sequences. *Proceedings of the conference. Association for Computational Linguistics. Meeting* **2019**, *2019*, 6558–6569.
 69. Wang, R.; Shivanna, R.; Cheng, D.Z.; Jain, S.; Lin, D.; Hong, L.; Chi, E.H. DCN V2: Improved Deep & Cross Network and Practical Lessons for Web-scale Learning to Rank Systems. *Proceedings of the Web Conference* **2021** **2020**.
 70. Wang, S.; McDermott, M.B.A.; Chauhan, G.; Hughes, M.C.; Naumann, T.; Ghassemi, M. MIMIC-Extract: a data extraction, preprocessing, and representation pipeline for MIMIC-III. *Proceedings of the ACM Conference on Health, Inference, and Learning* **2019**.
 71. Tang, S.; Davarmanesh, P.; Song, Y.; Koutra, D.; Sjoding, M.W.; Wiens, J. Democratizing EHR analyses with FIDDLE: a flexible data-driven preprocessing pipeline for structured clinical data. *Journal of the American Medical Informatics Association : JAMIA* **2020**, *27*, 1921 – 1934.
 72. Shang, J.; Xiao, C.; Ma, T.; Li, H.; Sun, J. GAMENet: Graph Augmented MEmory Networks for Recommending Medication Combination. *ArXiv* **2018**, *abs/1809.01852*.
 73. Choi, E.; Bahadori, M.T.; Schuetz, A.; Stewart, W.F.; Sun, J. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. *JMLR workshop and conference proceedings* **2015**, *56*, 301–318.
 74. Shin, B.; Hogan, J.; Adams, A.B.; Lynch, R.J.; Patzer, R.E.; Choi, J.D. Multimodal Ensemble Approach to Incorporate Various Types of Clinical Notes for Predicting Readmission. *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* **2019**, pp. 1–4.
 75. Yang, B.; Wu, L. How to leverage the multimodal EHR data for better medical prediction? In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2021.
 76. Qiao, Z.; Wu, X.; Ge, S.; Fan, W. MNN: Multimodal Attentional Neural Networks for Diagnosis Prediction. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2019.
 77. Park, J.; Cho, Y.; Lee, H.; Choo, J.; Choi, E. Knowledge Graph-based Question Answering with Electronic Health Records. In *Proceedings of the Machine Learning in Health Care*, 2020.
 78. Cai, B.; Xiang, Y.; Gao, L.; Zhang, H.; Li, Y.; Li, J. Temporal Knowledge Graph Completion: A Survey. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2022.
 79. Shi, L.; Li, S.; Yang, X.; Qi, J.; Pan, G.; Zhou, B. Semantic Health Knowledge Graph: Semantic Integration of Heterogeneous Medical Knowledge and Services. *BioMed Research International* **2017**, *2017*.
 80. Koshman, V.; Funkner, A.A.; Kovalchuk, S.V. An Unsupervised Approach to Structuring and Analyzing Repetitive Semantic Structures in Free Text of Electronic Medical Records. *Journal of Personalized Medicine* **2022**, *12*.
 81. Li, L.; Wang, P.; Yan, J.; Wang, Y.; Li, S.; Jiang, J.; Sun, Z.; Tang, B.; Chang, T.H.; Wang, S.; et al. Real-world data medical knowledge graph: construction and applications. *Artificial intelligence in medicine* **2020**, *103*, 101817.
 82. Hatef, E.; Rouhizadeh, M.; Tia, I.; Lasser, E.C.; Hill-Briggs, F.; Marsteller, J.; Kharrazi, H. Assessing the Availability of Data on Social and Behavioral Determinants in Structured and Unstructured Electronic Health Records: A Retrospective Analysis of a Multilevel Health Care System. *JMIR Medical Informatics* **2019**, *7*.

83. Liu, X.; Wang, H.; He, T.; Liao, Y.; Jian, C. Recent Advances in Representation Learning for Electronic Health Records: A Systematic Review. *Journal of Physics: Conference Series* **2022**, *2188*.
84. Zhang, Z.; Cao, L.; liang Chen, X.; Tang, W.; xiong Xu, Z.; Meng, Y. Representation Learning of Knowledge Graphs With Entity Attributes. *IEEE Access* **2020**, *8*, 7435–7441.
85. Tang, X.; Chen, L.; Cui, J.; Wei, B. Knowledge representation learning with entity descriptions, hierarchical types, and textual relations. *Inf. Process. Manag.* **2019**, *56*, 809–822.
86. Bordes, A.; Usunier, N.; García-Durán, A.; Weston, J.; Yakhnenko, O. Translating Embeddings for Modeling Multi-relational Data. In Proceedings of the Neural Information Processing Systems, 2013.
87. Wang, Z.; Zhang, J.; Feng, J.; Chen, Z. Knowledge Graph Embedding by Translating on Hyperplanes. In Proceedings of the AAAI Conference on Artificial Intelligence, 2014.
88. Lin, Y.; Liu, Z.; Sun, M.; Liu, Y.; Zhu, X. Learning Entity and Relation Embeddings for Knowledge Graph Completion. In Proceedings of the AAAI Conference on Artificial Intelligence, 2015.
89. Socher, R.; Chen, D.; Manning, C.D.; Ng, A. Reasoning With Neural Tensor Networks for Knowledge Base Completion. In Proceedings of the Neural Information Processing Systems, 2013.
90. Shi, B.; Weninger, T. Open-World Knowledge Graph Completion. In Proceedings of the AAAI Conference on Artificial Intelligence, 2017.
91. Lin, Y.; Liu, Z.; Sun, M. Knowledge Representation Learning with Entities, Attributes and Relations. In Proceedings of the International Joint Conference on Artificial Intelligence, 2016.
92. Tay, Y.; Luu, A.T.; Phan, M.C.; Hui, S.C. Multi-Task Neural Network for Non-discrete Attribute Prediction in Knowledge Graphs. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* **2017**.
93. Xie, R.; Liu, Z.; Jia, J.; Luan, H.; Sun, M. Representation Learning of Knowledge Graphs with Entity Descriptions. In Proceedings of the AAAI Conference on Artificial Intelligence, 2016.
94. Rasmussen, K.E.; Dolan, M.E. OncoCL: A Cancer Cell Ontology. In Proceedings of the International Conference on Biomedical Ontology, 2013.
95. Tao, C.; Sun, J.; Zheng, W.J.; Chen, J.; Xu, H. Drug Target Prediction for Colorectal Cancer by Combining Ontology and Network Approaches. In Proceedings of the International Conference on Biomedical Ontology, 2014.
96. Myneni, S.; Amith, M.T.; Geng, Y.; Tao, C. Towards an Ontology-driven Framework to Enable Development of Personalized mHealth Solutions for Cancer Survivors' Engagement in Healthy Living. *Studies in health technology and informatics* **2016**, *216*, 113 – 117.
97. Malt, A.M.; Jain, S.K.; Yang, P.C.; Harvey, K.; Warner, J.L. Computerized Approach to Creating a Systematic Ontology of Hematology/Oncology Regimens. *JCO clinical cancer informatics* **2018**, *2*.
98. Dinakarpanthian, D.; Liedtke, M.; Musen, M.A.; Dinakar, B. TOCSOC: A Temporal Ontology for Comparing the Survival Outcomes of Clinical Trials in Oncology. In Proceedings of the International Conference on Biomedical Ontology, 2018.
99. Cheng, B.; Zhang, J.; Liu, H.; Cai, M.; Wang, Y. Research on Medical Knowledge Graph for Stroke. *Journal of Healthcare Engineering* **2021**, *2021*.
100. Huang, J.; Lu, T.; Zhu, J.; Yu, W.; Zhang, T. Multi-relational knowledge graph completion method with local information fusion. *Applied Intelligence* **2021**, *52*, 7985 – 7994.
101. Bianchi, F.; Rossiello, G.; Costabello, L.; Palmonari, M.; Minervini, P. Knowledge Graph Embeddings and Explainable AI. In Proceedings of the Knowledge Graphs for eXplainable Artificial Intelligence, 2020.
102. Wang, Q.; Mao, Z.; Wang, B.; Guo, L. Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Transactions on Knowledge and Data Engineering* **2017**, *29*, 2724–2743.
103. Wang, Z.; Zhang, J.; Feng, J.; Chen, Z. Knowledge Graph Embedding by Translating on Hyperplanes. In Proceedings of the AAAI Conference on Artificial Intelligence, 2014.
104. Lin, Y.; Liu, Z.; Sun, M.; Liu, Y.; Zhu, X. Learning Entity and Relation Embeddings for Knowledge Graph Completion. In Proceedings of the AAAI Conference on Artificial Intelligence, 2015.
105. Li, L.; Wang, P.; Wang, Y.; Jiang, J.; Tang, B.; Yan, J.; Wang, S.; Liu, Y. PrTransH: Embedding Probabilistic Medical Knowledge from Real World EMR Data. *ArXiv* **2019**, *abs/1909.00672*.
106. Sun, Z.; Deng, Z.; Nie, J.Y.; Tang, J. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. *ArXiv* **2018**, *abs/1902.10197*.
107. Yang, B.; tau Yih, W.; He, X.; Gao, J.; Deng, L. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. *CoRR* **2014**, *abs/1412.6575*.

108. Nickel, M.; Tresp, V.; Kriegel, H.P. A Three-Way Model for Collective Learning on Multi-Relational Data. In Proceedings of the International Conference on Machine Learning, 2011.
109. Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, É.; Bouchard, G. Complex Embeddings for Simple Link Prediction. *ArXiv* **2016**, *abs/1606.06357*.
110. Nguyen, D.Q.; Nguyen, T.D.; Nguyen, D.Q.; Phung, D.Q. A Novel Embedding Model for Knowledge Base Completion Based on Convolutional Neural Network. In Proceedings of the North American Chapter of the Association for Computational Linguistics, 2017.
111. Dettmers, T.; Minervini, P.; Stenetorp, P.; Riedel, S. Convolutional 2D Knowledge Graph Embeddings. In Proceedings of the AAAI Conference on Artificial Intelligence, 2017.
112. Beltagy, I.; Lo, K.; Cohan, A. SciBERT: A Pretrained Language Model for Scientific Text. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2019.
113. Khattak, F.K.; Jebblee, S.; Pou-Prom, C.; Abdalla, M.; Meaney, C.; Rudzicz, F. A survey of word embeddings for clinical text. *Journal of biomedical informatics* **2019**, *100S*, 100057.
114. Alsentzer, E.; Murphy, J.R.; Boag, W.; Weng, W.H.; Jin, D.; Naumann, T.; McDermott, M.B.A. Publicly Available Clinical BERT Embeddings. *ArXiv* **2019**, *abs/1904.03323*.
115. Lee, J.; Yoon, W.; Kim, S.; Kim, D.; Kim, S.; So, C.H.; Kang, J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **2019**, *36*, 1234 – 1240.
116. Sang, S.; Yang, Z.; Liu, X.; Wang, L.; Lin, H.; Wang, J.; Dumontier, M. GrEDeL: A Knowledge Graph Embedding Based Method for Drug Discovery From Biomedical Literatures. *IEEE Access* **2019**, *7*, 8404–8415.
117. Bonner, S.; Barrett, I.P.; Ye, C.; Swiers, R.; Engkvist, O.; Hamilton, W.L. Understanding the Performance of Knowledge Graph Embeddings in Drug Discovery. *ArXiv* **2021**, *abs/2105.10488*.
118. Çelebi, 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 Bioinformatics* **2019**, *20*.
119. Crichton, G.K.O.; Guo, Y.; Pyysalo, S.; Korhonen, A. Neural networks for link prediction in realistic biomedical graphs: a multi-dimensional evaluation of graph embedding-based approaches. *BMC Bioinformatics* **2018**, *19*.
120. Grover, A.; Leskovec, J. node2vec: Scalable Feature Learning for Networks. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* **2016**.
121. Joshi, P.; Masilamani, V.; Mukherjee, A. A knowledge graph embedding based approach to predict the adverse drug reactions using a deep neural network. *Journal of biomedical informatics* **2022**, p. 104122.
122. Santos, A.; Colaço, A.R.; Nielsen, A.B.; Niu, L.; Strauss, M.T.; Geyer, P.E.; Coscia, F.; Albrechtsen, N.J.W.; Mundt, F.; Jensen, L.J.; et al. A knowledge graph to interpret clinical proteomics data. *Nature Biotechnology* **2022**, *40*, 692 – 702.
123. Agarwal, K.; Eftimov, T.; Addanki, R.; Choudhury, S.; Tamang, S.R.; Rallo, R. Snomed2Vec: Random Walk and Poincaré Embeddings of a Clinical Knowledge Base for Healthcare Analytics. *ArXiv* **2019**, *abs/1907.08650*.
124. 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.
125. Himmelstein, D.S.; Lizee, A.; Hessler, C.; Brueggeman, L.; Chen, S.L.; Hadley, D.; Green, A.J.; Khankhanian, P.; Baranzini, S.E. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife* **2017**, *6*.
126. Su, C.; Hou, Y.; Zhou, M.; Rajendran, S.; Maasch, J.R.M.A.; Abedi, Z.; Zhang, H.; Bai, Z.; Cuturrufo, A.; Guo, W.L.; et al. Biomedical discovery through the integrative biomedical knowledge hub (iBKH). *iScience* **2023**, *26* 4, 106460.
127. Youn, J.; Rai, N.; Tagkopoulos, I. Knowledge integration and decision support for accelerated discovery of antibiotic resistance genes. *Nature Communications* **2022**, *13*.
128. Suchanek, F.M.; Abiteboul, S.; Senellart, P. PARIS: Probabilistic Alignment of Relations, Instances, and Schema. *Proc. VLDB Endow.* **2011**, *5*, 157–168.
129. Maaroufi, M.; Choquet, R.; Landais, P.; Jaulent, M.C. Formalizing Mappings to Optimize Automated Schema Alignment: Application to Rare Diseases. *Studies in health technology and informatics* **2014**, *205*, 283–7.
130. Bachman, J.A.; Gyori, B.M.; Sorger, P.K. FamPlex: a resource for entity recognition and relationship resolution of human protein families and complexes in biomedical text mining. *BMC Bioinformatics* **2017**, *19*.
131. Hu, Y.; Nie, T.; Shen, D.; Kou, Y.; Yu, G. An integrated pipeline model for biomedical entity alignment. *Frontiers of Computer Science* **2021**, *15*, 1–15.

132. Ma, J.; Zhou, C.; Chen, Y.; Wang, Y.; Hu, G.; Qiao, Y. TeCre: A Novel Temporal Conflict Resolution Method Based on Temporal Knowledge Graph Embedding. *Inf.* **2023**, *14*, 155.
133. Faria, D.; Jiménez-Ruiz, E.; Pesquita, C.; Santos, E.; Couto, F.M. Towards Annotating Potential Incoherences in BioPortal Mappings. In Proceedings of the International Workshop on the Semantic Web, 2014.
134. He, Y.; Chen, J.; Dong, H.; Jiménez-Ruiz, E.; Hadian, A.; Horrocks, I. Machine Learning-Friendly Biomedical Datasets for Equivalence and Subsumption Ontology Matching **2022**. *abs/2205.03447*.
135. Chai, X. Diagnosis Method of Thyroid Disease Combining Knowledge Graph and Deep Learning. *IEEE Access* **2020**, *8*, 149787–149795.
136. Alawad, M.M.; Gao, S.; Shekar, M.C.; Hasan, S.M.S.; Christian, J.B.; Wu, X.C.; Durbin, E.B.; Doherty, J.A.; Stroup, A.M.; Coyle, L.; et al. Integration of Domain Knowledge using Medical Knowledge Graph Deep Learning for Cancer Phenotyping. *ArXiv* **2021**, *abs/2101.01337*.
137. Zhou, D.; Gan, Z.; Shi, X.; Patwari, A.; Rush, E.N.; Bonzel, C.L.; Panickan, V.A.; Hong, C.; Ho, Y.L.A.; Cai, T.A.; et al. Multiview Incomplete Knowledge Graph Integration with application to cross-institutional EHR data harmonization. *Journal of biomedical informatics* **2022**, p. 104147.
138. Al-Moslmi, T.; Ocaña, M.G.; Opdahl, A.L.; Veres, C. Named Entity Extraction for Knowledge Graphs: A Literature Overview. *IEEE Access* **2020**, *8*, 32862–32881.
139. Grishman, R.; Sundheim, B.M. Message Understanding Conference- 6: A Brief History. In Proceedings of the International Conference on Computational Linguistics, 1996.
140. Sharnagat, R. Named Entity Recognition: a Literature Survey. In Proceedings of the Indian Institute of Technology, 2014.
141. Yadav, V.; Bethard, S. A Survey on Recent Advances in Named Entity Recognition from Deep Learning models. *ArXiv* **2018**, *abs/1910.11470*.
142. Goyal, A.; Gupta, V.; Kumar, M. Recent Named Entity Recognition and Classification techniques: A systematic review. *Comput. Sci. Rev.* **2018**, *29*, 21–43.
143. Etzioni, O.; Cafarella, M.J.; Downey, D.; Popescu, A.M.; Shaked, T.; Soderland, S.; Weld, D.S.; Yates, A. Unsupervised named-entity extraction from the Web: An experimental study. *Artif. Intell.* **2005**, *165*, 91–134.
144. Zhang, S.; Elhadad, N. Unsupervised biomedical named entity recognition: Experiments with clinical and biological texts. *Journal of biomedical informatics* **2013**, *46*, 6, 1088–98.
145. Baevski, A.; Edunov, S.; Liu, Y.; Zettlemoyer, L.; Auli, M. Cloze-driven Pretraining of Self-attention Networks. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2019.
146. Peters, M.E.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L. Deep Contextualized Word Representations. *ArXiv* **2018**, *abs/1802.05365*.
147. Liu, L.; Ren, X.; Shang, J.; Peng, J.; Han, J. Efficient Contextualized Representation: Language Model Pruning for Sequence Labeling. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2018.
148. Liu, L.; Shang, J.; Xu, F.F.; Ren, X.; Gui, H.; Peng, J.; Han, J. Empower Sequence Labeling with Task-Aware Neural Language Model. In Proceedings of the AAAI Conference on Artificial Intelligence, 2017.
149. Akbik, A.; Blythe, D.A.J.; Vollgraf, R. Contextual String Embeddings for Sequence Labeling. In Proceedings of the International Conference on Computational Linguistics, 2018.
150. Li, J.; Sun, A.; Han, J.; Li, C. A Survey on Deep Learning for Named Entity Recognition. *IEEE Transactions on Knowledge and Data Engineering* **2018**, *34*, 50–70.
151. Shen, W.; Wang, J.; Han, J. Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions. *IEEE Transactions on Knowledge and Data Engineering* **2015**, *27*, 443–460.
152. Wang, C.; Sun, X.; Yu, H.; Zhang, W. Entity Disambiguation Leveraging Multi-Perspective Attention. *IEEE Access* **2019**, *7*, 113963–113974.
153. Daiber, J.; Jakob, M.; Hokamp, C.; Mendes, P.N. Improving efficiency and accuracy in multilingual entity extraction. In Proceedings of the International Conference on Semantic Systems, 2013.
154. Fang, W.; Zhang, J.; Wang, D.; Chen, Z.; Li, M. Entity Disambiguation by Knowledge and Text Jointly Embedding. In Proceedings of the Conference on Computational Natural Language Learning, 2016.
155. Zhu, G.; Iglesias, C.A. Exploiting semantic similarity for named entity disambiguation in knowledge graphs. *Expert Syst. Appl.* **2018**, *101*, 8–24.
156. Zhao, Q.; Hoi, S.C.H.; Liu, T.Y.; Bhowmick, S.S.; Lyu, M.R.; Ma, W.Y. Time-dependent semantic similarity measure of queries using historical click-through data. In Proceedings of the The Web Conference, 2006.
157. Moro, A.; Raganato, A.; Navigli, R. Entity Linking meets Word Sense Disambiguation: a Unified Approach. *Transactions of the Association for Computational Linguistics* **2014**, *2*, 231–244.

158. Usbeck, R.; Ngomo, A.C.N.; Röder, M.; Gerber, D.; Coelho, S.; Auer, S.; Both, A. AGDISTIS - Graph-Based Disambiguation of Named Entities Using Linked Data. In Proceedings of the International Workshop on the Semantic Web, 2014.
159. Piccinno, F.; Ferragina, P. From TagME to WAT: a new entity annotator. In Proceedings of the ERD '14, 2014.
160. Gong, L.; Yuan, Y.; Wei, Y.B.; Sun, X. A Hybrid Approach for Biomedical Entity Name Recognition. *2009 2nd International Conference on Biomedical Engineering and Informatics* **2009**, pp. 1–5.
161. Muzaffar, A.W.; Azam, F.; Qamar, U. A Relation Extraction Framework for Biomedical Text Using Hybrid Feature Set. *Computational and Mathematical Methods in Medicine* **2015**, 2015.
162. Munkhdalai, T.; Li, M.; Kim, T.; Namsrai, O.E.; Jeong, S.P.; Shin, J.; Ryu, K.H. Bio Named Entity Recognition Based on Co-training Algorithm. *2012 26th International Conference on Advanced Information Networking and Applications Workshops* **2012**, pp. 857–862.
163. Christen, P. Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection / Peter Christen. 2012.
164. Elmagarmid, A.K.; Ipeirotis, P.G.; Verykios, V.S. Duplicate Record Detection: A Survey. *IEEE Transactions on Knowledge and Data Engineering* **2007**, 19, 1–16.
165. Sarawagi, S.; Bhamidipaty, A. Interactive deduplication using active learning. *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* **2002**.
166. Pennington, J.; Socher, R.; Manning, C.D. GloVe: Global Vectors for Word Representation. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2014.
167. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed Representations of Words and Phrases and their Compositionality. In Proceedings of the Neural Information Processing Systems, 2013.
168. Bojanowski, P.; Grave, E.; Joulin, A.; Mikolov, T. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics* **2016**, 5, 135–146.
169. Peng, Y.; Yan, S.; Lu, Z. Transfer Learning in Biomedical Natural Language Processing: An Evaluation of BERT and ELMo on Ten Benchmarking Datasets. In Proceedings of the BioNLP@ACL, 2019.
170. Liu, F.; Shareghi, E.; Meng, Z.; Basaldella, M.; Collier, N. Self-Alignment Pretraining for Biomedical Entity Representations. In Proceedings of the North American Chapter of the Association for Computational Linguistics, 2020.
171. Michalopoulos, G.; Wang, Y.; Kaka, H.; Chen, H.H.; Wong, A. UmlsBERT: Clinical Domain Knowledge Augmentation of Contextual Embeddings Using the Unified Medical Language System Metathesaurus. In Proceedings of the North American Chapter of the Association for Computational Linguistics, 2020.
172. Singh A, Krishnamoorthy S, O.J. NeighBERT: Medical Entity Linking Using Relation Induced Dense Retrieval. In Proceedings of the Research Square, 2022.
173. Cai, X. A Novel Disease Diagnosis Method Using Combining Knowledge Graph and Deep Learning. *J. Medical Imaging Health Informatics* **2021**, 11, 1845–1851.
174. Agrawal, M.; Hegselmann, S.; Lang, H.; Kim, Y.; Sontag, D. Large Language Models are Few-Shot Clinical Information Extractors. In Proceedings of the Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP), 2022, pp. 1998–2022.
175. Chen, H.; Shen, X.; Lv, Q.; Wang, J.; Ni, X.; Ye, J. SAC-KG: Exploiting Large Language Models as Skilled Automatic Constructors for Domain Knowledge Graph. In Proceedings of the Proc. 62nd Annual Meeting of the Association for Computational Linguistics (ACL), 2024.
176. Lu, Y.; Goi, S.Y.; Zhao, X.; Wang, J. KARMA: Leveraging Multi-Agent LLMs for Automated Knowledge Graph Enrichment. *arXiv preprint arXiv:2502.06472* **2025**.
177. Yang, H.; Li, J.; Zhang, C.; Pazos Sierra, A.; Shen, B. Large Language Model-Driven Knowledge Graph Construction in Sepsis Care Using Multicenter Clinical Databases: Development and Usability Study. *Journal of Medical Internet Research* **2025**, 27, e65537. <https://doi.org/10.2196/65537>.
178. Sun, J.; Xu, C.; Tang, L.; Wang, S.; Lin, C.; Gong, Y.; Ni, L.M.; yeung Shum, H.; Guo, J. Think-on-Graph: Deep and Responsible Reasoning of Large Language Model on Knowledge Graph. In Proceedings of the International Conference on Learning Representations, 2023.
179. Edge, D.; Trinh, H.; Cheng, N.; Bradley, J.; Chao, A.; Mody, A.N.; Truitt, S.; Larson, J. From Local to Global: A Graph RAG Approach to Query-Focused Summarization. *ArXiv* **2024**, abs/2404.16130.
180. Gao, Y.; Li, R.; Caskey, J.R.; Dligach, D.; Miller, T.A.; Churpek, M.M.; Afshar, M. Leveraging A Medical Knowledge Graph into Large Language Models for Diagnosis Prediction. *ArXiv* **2023**, abs/2308.14321.
181. Xu, R.; Shi, W.; Yu, Y.; Zhuang, Y.; Jin, B.; Wang, M.D.; Ho, J.C.; Yang, C. RAM-EHR: Retrieval Augmentation Meets Clinical Predictions on Electronic Health Records. *ArXiv* **2024**, abs/2403.00815.

182. Xu, X.; others.; Miao, C. MedRAG: Enhancing Retrieval-Augmented Generation with Knowledge Graph-Elicited Reasoning for Healthcare Copilot. In Proceedings of the Proc. ACM Web Conference (WWW), 2025.
183. Wu, J.; Zhu, J.; Qi, Y.; Chen, J.; Xu, M.; Menolascina, F.; Jin, Y.; Grau, V. Medical Graph RAG: Evidence-Based Medical Large Language Model via Graph Retrieval-Augmented Generation. In Proceedings of the Proc. 63rd Annual Meeting of the Association for Computational Linguistics (ACL), 2025, pp. 28443–28467.
184. Rezaei, M.R.; Saadati Fard, R.; Parker, J.L.; Krishnan, R.G.; Lankarany, M. Agentic Medical Knowledge Graphs Enhance Medical Question Answering: Bridging the Gap Between LLMs and Evolving Medical Knowledge. In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP, 2025, pp. 12682–12701.
185. Jiang, X.; Zhang, R.; Xu, Y.; Qiu, R.; Fang, Y.; Wang, Z.; Tang, J.; Ding, H.; Chu, X.; Zhao, J.; et al. HyKGE: A Hypothesis Knowledge Graph Enhanced Framework for Accurate and Reliable Medical LLMs Responses. *arXiv preprint arXiv:2312.15883* **2024**.
186. Jiang, X.; Fang, Y.; Qiu, R.; Zhang, H.; Xu, Y.; Chen, H.; Zhang, W.; Zhang, R.; Fang, Y.; Chu, X.; et al. TC-RAG: Turing-Complete RAG's Case Study on Medical LLM Systems. *arXiv preprint arXiv:2408.09199* **2024**.
187. Chandak, P.; Huang, K.; Zitnik, M. Building a Knowledge Graph to Enable Precision Medicine. *Scientific Data* **2023**, *10*, 67.
188. Lewis, P.; Perez, E.; Piktus, A.; Petroni, F.; Karpukhin, V.; Goyal, N.; Kuttler, H.; Lewis, M.; tau Yih, W.; Rocktäschel, T.; et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *ArXiv* **2020**, *abs/2005.11401*.
189. Singhal, K.; Azizi, S.; Tu, T.; Mahdavi, S.; Wei, J.; Chung, H.W.; Scales, N.; Tanwani, A.K.; Cole-Lewis, H.J.; Pfohl, S.J.; et al. Large language models encode clinical knowledge. *Nature* **2022**, *620*, 172 – 180.
190. Li, J.; Cheng, X.; Zhao, W.X.; Nie, J.; rong Wen, J. HaluEval: A Large-Scale Hallucination Evaluation Benchmark for Large Language Models. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2023.
191. Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, Y.; Dai, Y.; Sun, J.; Guo, Q.; Wang, M.; et al. Retrieval-Augmented Generation for Large Language Models: A Survey. *ArXiv* **2023**, *abs/2312.10997*.
192. Chatteraj, S.; Walid, R.; Joshi, K.P. Semantically Rich Approach to Automating Regulations of Medical Devices. *2024 IEEE International Conference on Digital Health (ICDH)* **2024**, pp. 132–137.
193. Chatteraj, S.; Joshi, K.P. LLM Based Knowledge Graph Approach to Automating Medical Device Regulatory Compliance. *2025 IEEE International Conference on Big Data (BigData)* **2025**, pp. 321–328.
194. Sang, S.; Yang, Z.; Wang, L.; Liu, X.; Lin, H.; Wang, J. SemaTyP: a knowledge graph based literature mining method for drug discovery. *BMC Bioinformatics* **2018**, *19*.
195. Karim, M.R.; Cochez, M.; Jares, J.B.; Uddin, M.; Beyan, O.D.; Decker, S. Drug-Drug Interaction Prediction Based on Knowledge Graph Embeddings and Convolutional-LSTM Network. *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics* **2019**.
196. Ma, F.; You, Q.; Xiao, H.; Chitta, R.; Zhou, J.; Gao, J. KAME: Knowledge-based Attention Model for Diagnosis Prediction in Healthcare. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* **2018**.
197. Li, Y.; Qian, B.; Zhang, X.; Liu, H. Graph Neural Network-Based Diagnosis Prediction. *Big data* **2020**.
198. Li, W.; Li, H.; Yang, B.; Zhou, L.; Yang, X.; Zhang, M.; Wang, B. Knowledge-aware representation learning for diagnosis prediction. *Expert Systems* **2023**, *40*.
199. Sun, Z.; Dong, W.; Shi, J.; Huang, Z. Interpretable Disease Progression Prediction Based on Reinforcement Reasoning Over a Knowledge Graph. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **2023**.
200. Chen, D.; Ma, C.; Wu, Y. Clinical Knowledge Graph Embeddings with Hierarchical Structure for Thyroid Treatment Recommendation. *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCoM/CyberSciTech)* **2019**, pp. 432–436.
201. Zhang, Y.; Qian, S.; Fang, Q.; Xu, C. Multi-modal Knowledge-aware Hierarchical Attention Network for Explainable Medical Question Answering. *Proceedings of the 27th ACM International Conference on Multimedia* **2019**.
202. Pushpakom, S.P.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.J.; Guilleims, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery* **2018**, *18*, 41–58.

203. Saadat, H.; Shah, B.; Halim, Z.; Anwar, S. Knowledge graph-based convolutional network coupled with sentiment analysis towards enhanced drug recommendation. *IEEE/ACM transactions on computational biology and bioinformatics* **2022**, *PP*.
204. Walid, R.; Pande, K.; Elluri, L. Secure and Privacy-Compliant Data Sharing: An Essential Framework for Healthcare Organizations. 2024.
205. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic acids research* **1999**, *27* 1, 29–34.
206. Lin, X.; Quan, Z.; Wang, Z.J.; Ma, T.; Zeng, X. KGNN: Knowledge Graph Neural Network for Drug-Drug Interaction Prediction. In Proceedings of the International Joint Conference on Artificial Intelligence, 2020.
207. Gerl, P. Random Walks on Graphs. *Texts in Applied Mathematics* **2020**.
208. Stone, Z.; Shen, Y.; et al., C.J. Correction to: Optical variability of quasars with 20-year photometric light curves. *Monthly Notices of the Royal Astronomical Society* **2022**.
209. Eppig, J.T.; Smith, C.L.; Blake, J.A.; Ringwald, M.; Kadin, J.A.; Richardson, J.E.; Bult, C.J. Mouse Genome Informatics (MGI): Resources for Mining Mouse Genetic, Genomic, and Biological Data in Support of Primary and Translational Research. *Methods in molecular biology* **2017**, *1488*, 47–73.
210. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.L.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: tool for the unification of biology. *Nature Genetics* **2000**, *25*, 25–29.
211. Ardlie, K.G.; DeLuca, D.S.; Segrè, A.V.; et al., T.J.S. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **2015**, *348*, 648 – 660.
212. Huang, X.; Zhang, J.; Xu, Z.; Ou, L.; Tong, J. A knowledge graph based question answering method for medical domain. *PeerJ Computer Science* **2021**, *7*.
213. English, A.; Ford, C.A. The HIPAA privacy rule and adolescents: legal questions and clinical challenges. *Perspectives on sexual and reproductive health* **2004**, *36* 2, 80–6.
214. Sharma, S.K. Data Privacy and GDPR Handbook. 2019.
215. young Kim, D.; Elluri, L.; Joshi, K.P. Trusted Compliance Enforcement Framework for Sharing Health Big Data. *2021 IEEE International Conference on Big Data (Big Data)* **2021**, pp. 4715–4724.
216. Chattoraj, S.; Joshi, K.P. MedReg-KG: KnowledgeGraph for Streamlining Medical Device Regulatory Compliance. *2024 IEEE International Conference on Big Data (BigData)* **2024**, pp. 3382–3390.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.