

Identifying and Appraising Knowledge Graphs for Hospital Information Systems

Subhodip Mitra^{1,*}, Vikas Karamchand Dagar², Rashmi Ranjan Guru³, and Bhukya Subhash²

¹All India Institute of Medical Sciences, Mangalagiri, India

²All India Institute of Medical Sciences, Kalyani, India

³All India Institute of Medical Sciences, Jodhpur, India

*Corresponding author: subhodip123@gmail.com

Abstract. Hospital information systems (HIS) collect and store large amounts of clinical, administrative, and operational data from many sources. However, integrating these data across departments and making them interoperable remain difficult. Recent research has turned to knowledge graphs (KGs), which organize different types of healthcare data semantically and can support clinical decision-making, workflow optimization, and data exchange between systems. This article identifies and appraises existing KG applications in HIS on the basis of recent peer-reviewed literature. It surveys the implementation frameworks, ontological models, and integration architectures that researchers have adopted in hospitals, and it evaluates how well they address semantic heterogeneity, data silos, and real-time decision-making. The findings show that KGs have great potential in hospitals, but several barriers stand in the way of their adoption in the health information ecosystem, among them limited scalability, the lack of standardization of operational data, and low institutional readiness.

Keywords. Knowledge graphs; hospital information systems; clinical decision support; data interoperability; semantic web; healthcare ontologies

I. INTRODUCTION

Modern hospitals generate data from many different sources, including electronic health records (EHRs), laboratory information systems, radiology archives, pharmacy modules, and administrative databases. Despite the rapid growth of digital health tools over the past decade, most of this information stays locked inside departmental systems. The result is data silos that obstruct integrated patient care and make hospital management less efficient [1]. The challenge of bringing heterogeneous healthcare data together without losing semantic meaning has drawn researchers toward knowledge graphs (KGs) as a technology for hospital information systems (HIS).

A KG represents entities and their semantic relationships as a directed graph, in which nodes stand for real-world concepts and edges describe how those concepts relate to each other [2]. In healthcare, this means that clinical concepts such as diseases, symptoms, medications, procedures, and patient demographics can all be linked within a single searchable structure. Traditional relational databases struggle with the complexity and variability of medical data. KGs handle it better through flexible schemas and ontology-driven architectures that do not require a rigid, predefined structure [3].

KGs are relevant beyond the clinical setting as well. Hospitals also create and collect large amounts of structured and unstructured data through essential operating activities,

such as allocating resources, scheduling staff, managing the supply chain, and complying with regulations. Traditional database systems struggle to capture the relationships that matter across departments, workflows, and administration, whereas KGs are well suited to modeling such multi-relational structures [4].

This review surveys the current state of KG applications in HIS. It examines construction methodologies, integration frameworks, and downstream applications, and it assesses the challenges that stand in the way of broader institutional adoption. The literature reviewed was published between 2022 and 2025, and the selection focused on peer-reviewed studies that make empirical or systematic contributions to the field.

II. CONSTRUCTION OF HEALTHCARE KNOWLEDGE GRAPHS

Building a KG for a hospital setting typically requires a multi-stage pipeline of data acquisition, knowledge extraction, ontology mapping, and graph population (Fig. 1) [1]. The data fed into this pipeline come from structured EHR fields, semi-structured clinical notes, and unstructured free-text narratives written by clinicians. To convert unstructured clinical text into structured triples that a graph can store, researchers rely on natural language processing (NLP) techniques, especially named entity recognition (NER) and relation extraction [4]. Ontological frameworks play an important role here. Standard medical ontologies, including

TABLE 1. KEY ONTOLOGIES AND STANDARDS USED IN HOSPITAL KG CONSTRUCTION

Ontology or standard	Domain coverage	Role in hospital KGs
SNOMED CT	Clinical terminology	Entity standardization for diseases, procedures, and findings
ICD-10/11	Disease classification	Diagnostic coding and epidemiological mapping
LOINC	Laboratory observations	Standardization of laboratory test entities and clinical measurements
RxNorm	Medications	Drug entity normalization and interaction mapping
FHIR	Data exchange	Interoperability framework for cross-system KG integration
UMLS	Metathesaurus	Cross-ontology entity linking and semantic unification

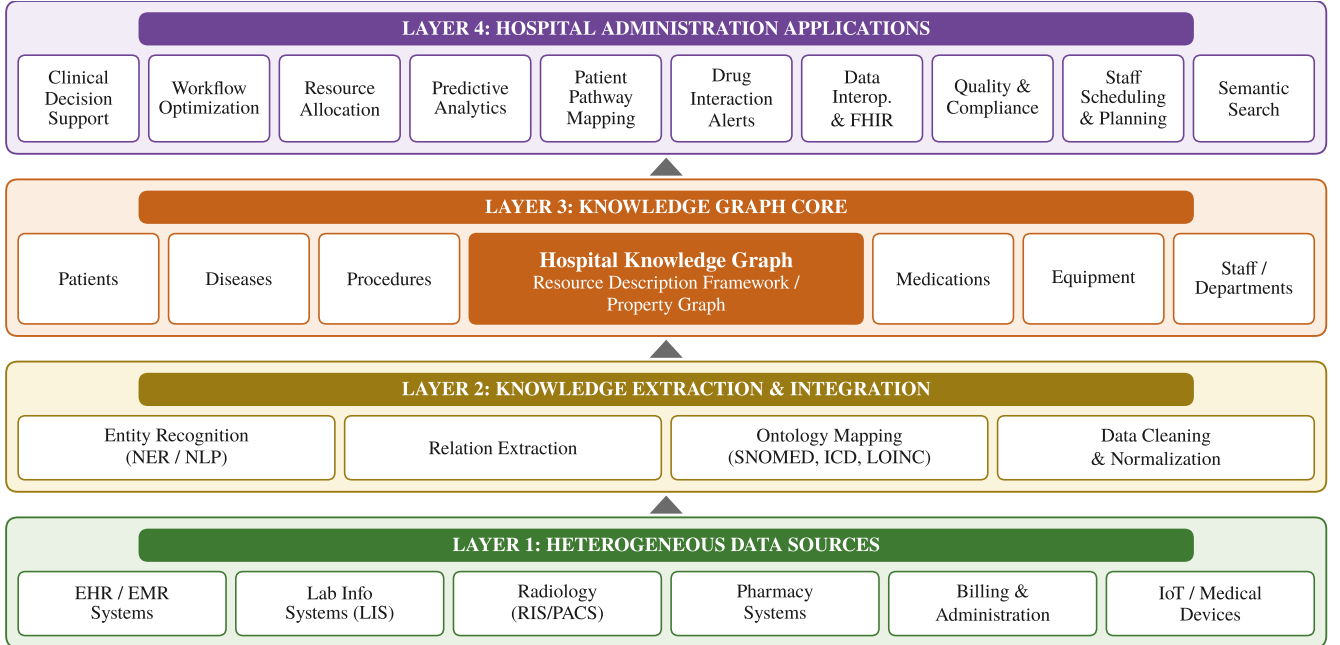


Fig. 1. Conceptual architecture of a knowledge graph-based hospital information system.

SNOMED CT, ICD-10/11, LOINC, and RxNorm, provide the underlying vocabulary that links each graph entity to a global terminology (Table 1) [3], [4].

According to Abu-Salih and colleagues [1], the most accurate construction methods for biomedical entity linking are hybrid approaches that combine rule-based extraction with machine learning classifiers. More recently, large language models (LLMs) have begun to change the way KGs are constructed. Yang and colleagues [5] reported that GPT-4-driven pipelines in healthcare settings can considerably accelerate construction while reducing manual expert effort.

III. APPLICATIONS IN HOSPITAL INFORMATION SYSTEMS

A. Clinical Decision Support

KGs support clinical decisions in hospitals through semantic reasoning over connected medical entities (Table 2). Shang and colleagues [6] built an EHR-oriented KG system that performs collaborative reasoning across multiple hospitals and keeps patient data private through blockchain-based synchronization. Their system improved alerting for the progression of chronic kidney disease, especially in departments outside nephrology, where such progression might otherwise be missed. The DR.KNOWS system developed by Gao and colleagues [7] takes a different route: it feeds UMLS-based KG information into an LLM to predict

diagnoses from EHR data. The combined KG-LLM system improved diagnostic accuracy compared with LLM-only approaches.

B. Data Interoperability and Integration

Passing data between hospital systems in a semantically meaningful way remains a major unsolved problem. Data models are frequently incompatible, and context is lost in translation. Sachdeva and Bhalla [2] explored whether KG structures based on openEHR archetypes and OWL ontologies can support semantic data exchange between EHR systems. In an independent study, Tabari and colleagues [8] examined implementations of FHIR-based data models. They found that KGs created from OMOP databases with FHIR-Ontop systems can form a virtual KG that connects otherwise incompatible repositories of clinical data. The growing convergence of FHIR standards and KG architectures appears to offer a realistic path to semantic interoperability across multi-hospital networks.

C. Patient-Centric Knowledge Representation

Al Khatib and colleagues [3] examined KGs built around a patient-centered representation. Patient-centric knowledge graphs (PCKGs) differ from other data structures in that they organize information around the patient instead of the disease. They unify the patient's longitudinal health records,

TABLE 2. SUMMARY OF KEY KNOWLEDGE GRAPH STUDIES IN HOSPITAL INFORMATION SYSTEMS

Study	Focus area	Key contribution	Year
Abu-Salih [1]	KG construction	Taxonomy of healthcare KG methods	2023
Sachdeva and Bhalla [2]	Interoperability	KG for semantic EHR exchange	2022
Al Khatib [3]	Patient-centric KG	Survey of PCKG methods	2024
Yang [5]	LLM-driven KG	GPT-4 sepsis KG construction	2025
Shang [6]	Decision support	Multi-hospital collaborative KG	2024
Gao [7]	LLM and KG	KG-enhanced diagnosis prediction	2025
Tabari [8]	FHIR integration	Review of FHIR data models	2024
Cui [9]	Review	Healthcare KG resources and applications	2025

genomics, social determinants, and treatment history in one representation. The authors caution that unless these ontologies are designed carefully with standards such as SNOMED CT and FHIR, the resulting graphs will be inconsistent and will not be interoperable across institutions.

D. Hospital Workflow and Operations

KGs are also beginning to show their value outside the clinic, in hospital operations. When departmental processes, staffing patterns, equipment use, and patient flow data are linked semantically in a graph, process mining and workflow optimization become possible at an institutional scale. The reasoning abilities of KGs can help hospitals identify bottlenecks, forecast resource needs, and improve scheduling by adding semantic inference over linked operational data [9]. Temporal reasoning can be added to these graphs to support longitudinal analysis of operational trends and predictive capacity planning. Fig. 2 gives a thematic overview of these application domains together with the key challenges discussed in Section IV.

IV. CHALLENGES AND BARRIERS TO ADOPTION

Despite the progress described above, several hurdles make KGs difficult to deploy. Scalability is the most pressing. According to Hofer and colleagues [10], a KG that spans several hospitals must store millions of entities and relationships while still answering queries in real time. They also draw attention to three persistent issues that cut across the field: (1) scalability combined with data sparsity; (2) heavy reliance on domain experts for ontology design; and (3) fragmentation across the construction pipeline, where errors at one stage are compounded at the next [10].

Standardization is another weak link. Although ontologies such as SNOMED CT and ICD cover clinical terminology well, there is no commonly accepted schema for hospital operational data such as staffing, scheduling, and supply chain records. This gap makes it difficult to construct KGs that work across institutions [9]. Data quality is also a serious concern. EHR-based KGs inherit most of the inconsistencies found in hospital EHRs, including variation in how clinicians document findings, missing values, and coding errors [4].

Privacy and security add a further complication. Put simply, an individual patient's record cannot be used to build a KG without that patient's consent. Shang and colleagues [6] used blockchain-based synchronization to address this problem in their multi-hospital system, but such methods incur computational overhead. Institutional readiness is a further question. A KG system cannot realistically be applied in a

hospital unless the technical infrastructure is developed and operational, staff are trained, and organizational change management is in place, and these requirements are often overlooked [9]. Healthcare data also change constantly: new patients arrive, diagnoses evolve, and treatment protocols change. Keeping a KG up to date therefore requires continuous maintenance, which at present is mostly manual and resource-intensive [10]. The situation is made worse by the absence of automated tools for verifying the completeness and correctness of a clinical KG.

V. FUTURE DIRECTIONS

The most active area of development is currently the integration of LLMs with KGs. Recent research shows that LLMs can automate entity extraction, relation extraction, and even parts of ontology construction, which substantially reduces the expert labor these tasks normally require [5]. This combination is relevant to hospitals because much of their data exists as unstructured clinical narratives. Since KGs make reasoning more structured and LLMs handle natural language well, the two can work together on data that neither handles well alone [9].

Federated KG architectures are another direction worth watching [6]. In a federated setting, each hospital keeps its own local KG while still being able to reason collaboratively with other hospitals. Privacy-preserving mechanisms such as differential privacy and secure multi-party computation make this possible. This matters in healthcare, where regulatory requirements and institutional policies closely restrict what data can be shared and how.

The development of FHIR interoperability standards offers a viable way to bring KG functions into existing hospital IT systems [8]. As FHIR adoption spreads globally, there is a considerable opportunity to establish FHIR-native KG construction pipelines. Such tools could generate semantic graph structures automatically from standardized clinical resources. For hospitals with FHIR-compliant infrastructure, these pipelines would substantially lower the technical barrier to deploying KGs.

KGs that model administrative workflows such as procurement, credentialing, regulatory reporting, maintenance, and supply chain management would be practically useful, but they have not yet received attention from hospital administrators. By merging clinical and administrative knowledge in a single graph architecture, hospitals could make operational decisions informed by clinical context, and the reverse. There are also no domain-specific benchmarks for evaluating KGs in hospital administration. The metrics in current use are largely adapted from general biomedical

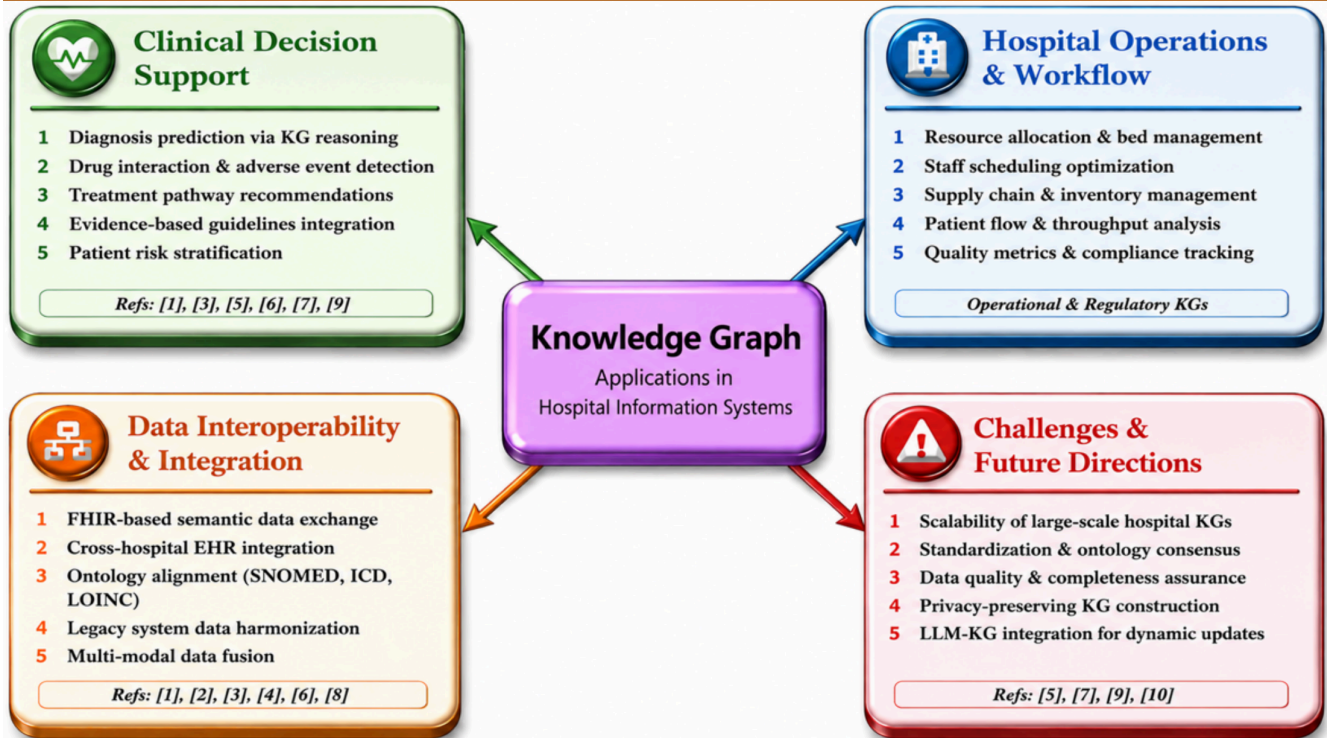


Fig. 2. Thematic categorization of knowledge graph applications in hospital information systems.

informatics and may not capture what matters in an operational environment [10].

Temporal KGs, which track how medical knowledge and patient states change over time, are also clearly relevant to hospital operations. A temporal KG can help monitor patients over time, forecast resource needs, and adapt clinical pathways as a patient's condition changes. Combining temporal reasoning with streaming data from Internet of Things (IoT) devices in the hospital could lead to systems that anticipate needs instead of merely reacting to them.

VI. CONCLUSION

Hospitals can use KGs to build precise representation models through which the health sector as a whole can be assessed. This review notes the growing use of KGs for clinical decision support, data interoperability, and patient-centered care, along with the existing methodological frameworks for building these graphs. The layered architecture described above, which spans data sources, knowledge extraction, graph representation, and applications, provides a useful model for seeing how KGs fit into a hospital.

The challenges, however, are real. Adoption faces constraints of scalability, standardization, data quality, and institutional readiness, and these issues are not easily fixed. The combination of KGs with LLMs and FHIR standards is a promising step in the right direction. Still, realizing that promise in working hospital systems will require sustained collaboration among hospital administrators, clinical informaticians, and data engineers. For administrators especially, KGs are not merely a technology upgrade. They represent a shift toward a management strategy in which clinical, operational, and administrative data are interconnected through a common indexable structure, and in which decisions in one domain can draw on knowledge from the others.

ACKNOWLEDGMENT

The authors gratefully acknowledge the faculty and research community for their contributions and support in the preparation of this manuscript.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and preparation of this work and approved the final version of the manuscript.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

DECLARATION OF GENERATIVE AI USE

The authors declare that no generative artificial intelligence tools were used in the preparation of this manuscript. The authors take full responsibility for the content of the publication.

DATA AVAILABILITY

No new data were created or analyzed in this study.

ETHICS STATEMENT

This study is a review of published literature and did not involve human participants or animals; ethical approval was not required.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- [1] Bilal Abu-Salih, Muhammad AL-Qurishi, Mohammed Alweshah, Mohammad AL-Smadi, Reem Alfayez, and Heba Saadeh, "Healthcare knowledge graph construction: A systematic review of the state-of-the-art, open issues, and opportunities," *J. Big Data*, vol. 10, no. 1, Art. no. 81, 2023, doi: 10.1186/s40537-023-00774-9.

- [2] Shelly Sachdeva and Subhash Bhalla, "Using knowledge graph structures for semantic interoperability in electronic health records data exchanges," *Information*, vol. 13, no. 2, Art. no. 52, 2022, doi: 10.3390/info13020052.
- [3] Hassan S. Al Khatib, Subash Neupane, Harish Kumar Manchukonda, Noorbakhsh Amiri Golilarz, Sudip Mittal, Amin Amirlatifi, and Shahram Rahimi, "Patient-centric knowledge graphs: A survey of current methods, challenges, and applications," *Front. Artif. Intell.*, vol. 7, Art. no. 1388479, 2024, doi: 10.3389/frai.2024.1388479.
- [4] Lino Murali, G. Gopakumar, Daleesha M. Viswanathan, and Prema Nedungadi, "Towards electronic health record-based medical knowledge graph construction, completion, and applications: A literature study," *J. Biomed. Inform.*, vol. 143, Art. no. 104403, 2023, doi: 10.1016/j.jbi.2023.104403.
- [5] Hao Yang, Jiaxi Li, Chi Zhang, Alejandro Pazos Sierra, and Bairong Shen, "Large language model-driven knowledge graph construction in sepsis care using multicenter clinical databases: Development and usability study," *J. Med. Internet Res.*, vol. 27, Art. no. e65537, 2025, doi: 10.2196/65537.
- [6] Yong Shang, Yu Tian, Kewei Lyu, Tianshu Zhou, Ping Zhang, Jianghua Chen, and Jingsong Li, "Electronic health record-oriented knowledge graph system for collaborative clinical decision support using multicenter fragmented medical data: Design and application study," *J. Med. Internet Res.*, vol. 26, Art. no. e54263, 2024, doi: 10.2196/54263.
- [7] Yanjun Gao, Ruizhe Li, Emma Croxford, John Caskey, Brian W. Patterson, Matthew Churpek, Timothy Miller, Dmitriy Dligach, and Majid Afshar, "Leveraging medical knowledge graphs into large language models for diagnosis prediction: Design and application study," *JMIR AI*, vol. 4, Art. no. e58670, 2025, doi: 10.2196/58670.
- [8] Parinaz Tabari, Gennaro Costagliola, Mattia De Rosa, and Martin Boeker, "State-of-the-art Fast Healthcare Interoperability Resources (FHIR)-based data model and structure implementations: Systematic scoping review," *JMIR Med. Inform.*, vol. 12, Art. no. e58445, 2024, doi: 10.2196/58445.
- [9] Hejie Cui, Jiaying Lu, Ran Xu, Shiyu Wang, Wenjing Ma, Yue Yu, Shaojun Yu, Xuan Kan, Chen Ling, Liang Zhao, Zhaohui S. Qin, Joyce C. Ho, Tianfan Fu, Jing Ma, Mengdi Huai, Fei Wang, and Carl Yang, "A review on knowledge graphs for healthcare: Resources, applications, and promises," *J. Biomed. Inform.*, vol. 169, Art. no. 104861, 2025, doi: 10.1016/j.jbi.2025.104861.
- [10] Marvin Hofer, Daniel Obraczka, Alieh Saeedi, Hanna Köpcke, and Erhard Rahm, "Construction of knowledge graphs: Current state and challenges," *Information*, vol. 15, no. 8, Art. no. 509, 2024, doi: 10.3390/info15080509.