A Medical Knowledge Management Mechanism with Knowledge Hypergraph Theory

Qi Wang
College of Science and Technology
Wuhan University of Science and Technology
Wuhan, Hubei, China
1024574503@qq.com

Qiyuan Li
College of Science and Technology
Wuhan University of Science and Technology
Wuhan, Hubei, China
vickyuan@wust.edu.cn

Chao Gao
College of Science and Technology
Wuhan University of Science and Technology
Wuhan, Hubei, China
gchao@ieee.org

Haijiang Liu
College of Science and Technology
Wuhan University of Science and Technology
Wuhan, Hubei, China
alecliu@ontoweb.wust.edu.cn

Fangfang Xu*
College of Science and Technology
Wuhan University of Science and Technology
Wuhan, Hubei, China
xuff@wust.edu.cn

Jinguang Gu
College of Science and Technology
Wuhan University of Science and Technology
Wuhan, Hubei, China
simon@wust.edu.cn

Abstract

The construction of medical knowledge graphs(KGs) structures complex medical data to support precise clinical decision-making and knowledge discovery. A core challenge in building medical KGs is effectively representing and processing complex clinical data. Existing triple-based medical knowledge representation methods often fail to fully capture and express the complexity of the data, particularly when dealing with higher-order relationships involving multiple entities. Additionally, the current "Attribute-Concept-Event" three-layer knowledge organization framework shows limitations in representing procedural knowledge in medical scenarios, falling short of the hierarchical and dynamic requirements of evidence-based medicine. To address these issues, we propose a four-layer medical knowledge organization method based on hypergraph theory. By incorporating hypergraph theory and a narrative layer into the KG, we provide a more flexible and dynamic framework for representing complex clinical information. Based on this approach, we construct a four-layer KG, RJUA-HKG, using real QA clinical data from urology as an example. Experiments show that, compared to traditional KGs, RJUA-HKG significantly reduces relational complexity and retrieval time while accurately capturing changes in diagnostic and treatment procedures.

CCS Concepts

• **Applied computing** → Life and medical sciences; Health care information systems.

*Corresponding author.

ISAIMS 2024, August 13–17, 2024, Amsterdam, Netherlands

ACM ISBN 979-8-4007-1782-6/24/08
https://doi.org/10.1145/3706890.3706998

Keywords

Knowledge organization, Knowledge graph, Knowledge hypergraph theory, Medical knowledge management mechanism

ACM Reference Format:

Qi Wang, Qiyuan Li, Chao Gao, Haijiang Liu, Fangfang Xu, and Jinguang Gu. 2024. A Medical Knowledge Management Mechanism with Knowledge Hypergraph Theory. In *2024 5th International Symposium on Artificial Intelligence for Medicine Science (ISAIMS 2024), August 13–17, 2024, Amsterdam, Netherlands.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3706890.3706998

1 Introduction

As a structured knowledge representation method, knowledge graphs(KGs) have been widely applied in vertical domains such as education[1], financial services[2], and e-commerce[3]. In healthcare, the rapid adoption of healthcare information systems and the expansion of the B2C healthcare model[4] have generated a vast amount of medical data[5]. Researchers have constructed numerous KGs[6-7] from this data to support various clinical applications.

Although KGs have strong modeling capabilities for entity relationships, they still face challenges when dealing with complex and multi-dimensional relationships[8]. In recent years, medical KGs have primarily been constructed using triplets[9]. However, these methods often oversimplify the complexity of the data stored within KGs[10]. This is particularly evident in the case of hyper-relational data that connects more than two entities, where the loss of high-order structural information becomes more severe when representing complex medical concepts and relationships[11]. For example, the relationships between diseases, symptoms, medications, and treatments may involve multiple attributes and conditions, requiring additional contextual information for accurate representation. As shown in Figure 1, more than 33.3%[12] of entities and 61%[13] of relationships in the Freebase KG cannot be represented by binary relations.

Medical KGs usually focus on binary relationships, using triplets to represent them. This approach is not concise enough for multi-dimensional relationships, such as between multiple diseases and

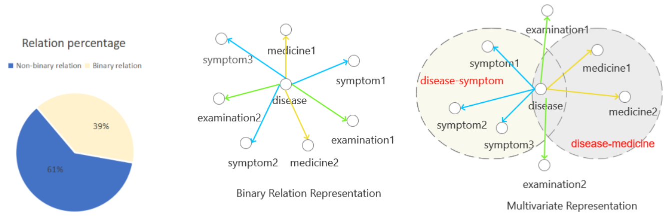

Figure 1: Many relationships are non-binary, but existing medical knowledge graphs cannot express these. Similar relationships are often represented in an inefficient way.

symptoms. This leads to similar relationships being represented more than once (e.g. "treatable by," "has symptom"). Furthermore, medical KGs typically use a three-layer structure of "Attribute-Concept-Event," which fails to represent procedural knowledge in medical scenarios, such as diagnosis and prognosis. Therefore, when constructing medical KG ontologies, it is crucial to ensure the simplicity of multi-dimensional relationships while also meeting the need for evidence-based and interpretable knowledge.

A hypergraph can connect any number of vertices[14], making it a graph structure that can represent multi-dimensional relationships and connections between nodes. Hypergraph models are being used in many areas, including traffic prediction[15], recommendation systems[16], and financial risk control[17]. For instance, Gao et al.[18] developed a model to help with knowledge retrieval which applies hypergraph theory to improve ontology construction.

To meet the needs of evidence-based medicine, existing medical datasets often capture procedural knowledge in unstructured formats, such as conversations. For example, RJUA-QA[1] uses doctor-patient dialogue data to represent urological diagnosis and treatment processes. To accurately represent procedural transitions in medical scenarios within a KG, we expand the traditional three-layer ontology structure to a four-layer framework: "Attribute-Concept-Event-Narrative." This new structure organizes procedural knowledge in a structured, searchable manner, enhancing the interpretability and reliability of the graph.

In conclusion, we propose a four-layer medical knowledge organization method based on hypergraph theory. Integrating hypergraph theory and a narrative layer to the KG makes it easier to represent complex clinical information and use KGs in clinical applications. Based on this, we also use a real urological QA dataset to create a four-layer hypergraph-based KG called RJUA-HKG. RJUA-HKG is simpler and faster than traditional KGs. It also shows changes in diagnosis and treatment.

Our contributions are as followings:

1. We propose a knowledge graph construction method based on hypergraph theory. It is better than the traditional triple-based knowledge representation and can represent complex relationships in more ways.

2. We introduce a four-layer medical knowledge organization method. We add a narrative layer to the existing structure, allowing the KG to capture the nature of the diagnosis and treatment process.

3. Use a real urological QA dataset to construct a four-layer KG, RJUA-HKG, based on hypergraph theory. This reduces relationship complexity and retrieval time while accurately describing procedural changes in diagnosis and treatment. This provides clinicians with evidence-based knowledge and improves retrieval efficiency.

2 Related Work

2.1 Traditional medical KG construction

As a kind of semantic network[19], a KG contains entities and the relationships between them. Medical KGs are a current research focus, where the main processes include medical ontology construction, knowledge extraction and medical knowledge fusion. Technically speaking, the main objective of medical knowledge extraction is to extract medical entities, relationships, and attributes from complex data sources. Traditional knowledge extraction methods mostly focus on individual named entity recognition(NER)[20] or binary relationship extraction(RE)[21] tasks. With the development of large language models(LLMs), models such as FSUIE[22] and InstructIE[23] have been proposed for information extraction(IE) tasks. However, due to the complexity of the medical field, when processing textual data, a certain complication mentioned in one sentence may become the disease itself introduced in the next sentence, and a disease name may be mentioned in a training text segment with multiple identities. These models are not conducive to IE. Therefore, Ant and Zhejiang University jointly designed the OneKE[2] knowledge extraction framework based on LLAMA2[24] and improved the model's reasoning ability by using Group-Query Attention. This approach has resolved the performance degradation issues caused by semantic confusion and differences in pattern querying methods, enhancing the generalization capability of LLMs in structured IE.

Medical KG research is growing swiftly. Despite the triple representation's popularity, it struggles with complex reasoning due to high computational demands, hindering efficiency on large graphs. The intricate and ambiguous 1-N and N-N relationships further complicate matters. Hypergraphs overcome these challenges by representing complex, multi-relational structures more efficiently, enhancing knowledge representation.

2.2 Hypergraph-based KG construction

Hypergraph is a graph structure that represents multi-dimensional relationships. It captures complex relationships and structures, and preserves relational attributes between entities, enabling more efficient knowledge representation. A hypergraph[25] is a graph where a hyperedge connects three or more vertices. C. Berge first introduced the concept of a hypergraph in 1970. It is a way of representing complex relationships. With the development of internet technologies, hypergraph theory has been widely applied in various fields. In clinical cases, a patient's medical history, symptoms, diagnoses and treatments are often linked and changing. Traditional knowledge organization structures struggle to capture this complexity, limiting the expressiveness and flexibility of KGs in real-world applications. It is difficult to organize medical knowledge over time.

[1]http://data.openkg.cn/dataset/rjua-qadatasets

[2]https://github.com/zjunlp/DeepKE/blob/main/example/llm/OneKE.md

This makes it hard to model key characteristics in patient information. We introduce a new way to organize knowledge. It adds a narrative layer on top of the traditional three-layer structure. This layer connects information in a more natural way. The narrative layer adds context to the KG and shows how diagnoses and treatments are structured. The four-layer medical knowledge method based on hypergraph theory expands the use of KGs and offers new ways to manage medical knowledge. It helps information to be integrated more efficiently and makes clinical decisions easier.

3 Preliminary

3.1 *Hypergraph definition*

A hypergraph is a specialized structure in graph theory, formally defined as $H = (V, E)$, where V is the set of vertices, and E is the set of edges. Specifically, E is a family of subsets of V , denoted as $E = \{(e_i)\}_{i \in I}$, where I is a finite index set. Each e_i is a subset of V, representing a hyperedge. Sometimes, $V(H)$ is used to denote the vertex set of hypergraph H, and $E(H)$ is its edge set. This structure allows hypergraphs to more flexibly represent complex relationships, surpassing the limitation in traditional graphs where edges can only connect two vertices.

3.2 *Three-layer knowledge organization*

The three-layer knowledge organization framework consists of three main core layers:

1. **Attribute Layer:** This layer describes an entity's characteristics. It includes basic information like name and category, as well as numbers like quantity. This layer provides a complete and accurate picture of the entity's characteristics.

2. **Concept Layer:** The conceptual layer is about creating categories and showing relationships. This layer shows how things are connected and provides a way to organize knowledge.

3. **Event Layer:** The event layer records how things change and happen in a system. It tracks changes in state and events involving entities. The event layer also shows when events happen and what causes them.

This three-layer structure shows how to organize knowledge. It will help people to find, use, and manage knowledge more effectively.

4 Method

4.1 *The four-layer organization of medical knowledge*

To effectively capture the hierarchical and dynamic aspects of the diagnosis and treatment process, we propose a four-layer knowledge organization framework: Attribute-Concept-Event-Narrative. As shown in Table 1.

- **Attribute Layer**: Primarily describes disease-related entities using data attributes.
- **Concept Layer**: Focuses on modeling relationships between medical concepts.
- **Event Layer**: Describe medical diagnostic events and support the representation of event relationships.
- **Narrative Layer**: The narrative layer organizes medical knowledge in a goal-driven manner, modeling features such as traceability of patient treatment information to provide comprehensive support for downstream tasks.

Figure 2 shows how the framework links key knowledge and causes in the medical process and shows how diagnosis and treatment work together. This helps us understand how to manage disease and make clinical decisions.

4.2 Hypergraph-based KG construction

This section looks at how KGs are built using hypergraph theory. This process includes steps like building an ontology, getting knowledge, and combining it. It aims to provide new ways to handle complex data and find knowledge.

4.2.1 Ontology construction. ***Ontology:*** It's a way of representing knowledge. In medical KGs, ontology construction aims to create a clear and formalized conceptual system. This means defining what an entity is, how they relate to each other, and what they are like. Medical concepts are represented as nodes in the hypergraph, using medical terminology standards for hierarchical and categorical modeling. As illustrated in Figure 3, this framework covers diseases, symptoms, causes, medications, treatments, and more. It supports medical research and clinical practice and helps with data analysis, decision support, and new knowledge.

Hyperedges: KG relationships are many-to-many with hyperedges representing connections. In a hypergraph, nodes represent diseases, symptoms, and treatments. Hyperedges link nodes. We can create a hyperedge called "Disease-Medication Association" to link diseases with treatments. We can also link diseases with the same symptoms or causes. Hypergraphs can handle more complex relationships. A medication may treat a disease by affecting several processes. A hyperedge could include the medication, processes, and disease, showing how the medication affects the disease.

Attributes: provide detailed features and information about each entity. For example, in the medical domain, the attributes of a disease may include "Disease Description," "Disease Classification," and "ICD Code"; whereas the attributes of a medication may cover "Side Effects," "Recommended Dosage," and "Interactions."

We define the possible sets of attributes for the aforementioned concept types in Table 2.

4.2.2 Knowledge acquisition. In the construction of medical KGs, the knowledge acquisition process entails converting knowledge from medical literature or clinical workflows into structured representations. It can be divided into three main steps:

- **Determining Data Sources:** The first step is to identify relevant data sources. Typically, this includes specialized biomedical databases, scientific literature databases, and various online medical consultation platforms.
- **Knowledge Extraction:** Automated knowledge extraction methods include NER and RE. Effective extraction strategies include rule-based, pattern-based and deep learning. Deep learning is the most popular and effective method. They use models like BERT and GPT for specific data which can learn a lot during pre-training. To improve the model, data can be augmented, tasks can be learned at once, and it can be trained to resist attacks.

Table 1: Defines the four-layer medical knowledge organization framework.

Knowledge Framework	Ontology	
	Entity type	Relationship type
Narrative Layer	(1) Entities involved in the “complaint/symptom-examination-diagnosis-treatment” process; (2) Rules of evidentiary reasoning;	(1) Causal relationships (causes, predisposes, produce, etc.); (2) Evolutionary relationships; (3) Knowledge source relationships;
Event Layer	(1) Medical events as defined in the medical literature and clinical guidelines; (2) Medical events that occur objectively in the clinic (diagnosis, intervention, follow-up, efficacy prediction, etc.); (3) Events described in the chief complaint (anxiety, insomnia, etc.); (4) Other events;	(1) Associative relationships between events; (2) Attributes of events such as time, place, and people; (3) Event calculation rules;
Concept Layer	(1) Entities related to UMLS, SNOMED-CT, and ICD-11 diagnostic systems, medicines, surgical procedures, and medical clinical datasets; (2) Characteristic conceptual entities formed by specific data; (3) Other entities;	Literature and Clinical Entity Relationships
Attribute Layer	(1) Medical ontology attributes: e.g., those defined by UMLS, SNOMED-CT, DSM-5, and ICD-11; (2) Characteristic attributes included in specific data; (3) Other attributes;	

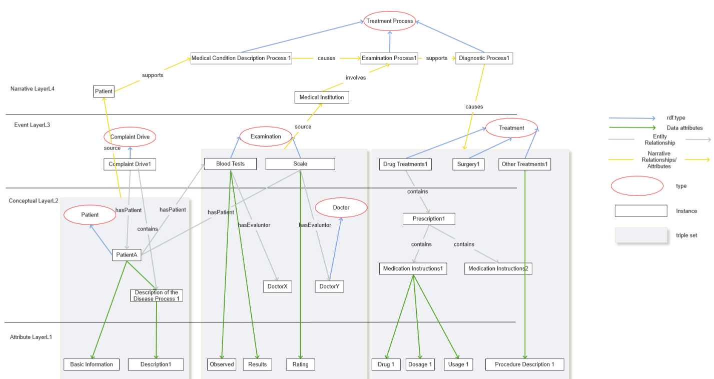

Figure 2: An example of four-layer medical knowledge organization.

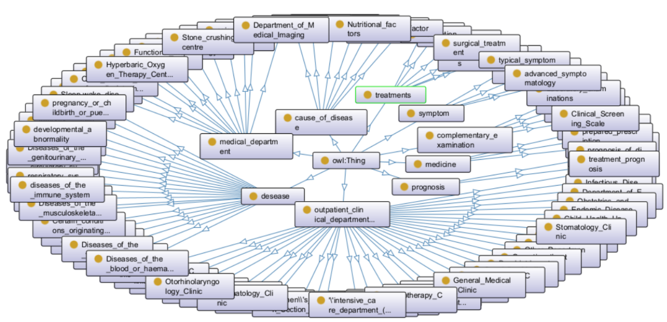

Figure 3: Ontology hierarchy classification diagram.

Table 2: Medical Field Attribute Set.

Concept	Attribute
Diseases	Disease Name/Disease Description/Disease Classification/ICD Code/Morbidity/Mortality /Complications/SNOMED-CT codes/Prognosis/Related Genes/Associated Drugs/Severity
Symptomatic	Symptom Name/Symptom Description/Related Diseases/Related Tests/ Related Drugs/Duration/Frequency of Symptoms
Etiology	Name of the cause/Cause Description/Associated Disease/Risk factors for Etiology
Medicines	Drug Name/Drug Classification/Dosage/Drug Usage/Side effects/Contraindications/Drug Interactions/Drug Price/Manufacturers/Drug Approval Number/Efficacy/Safety
Treatments	Treatment/Treatment Plan/Effectiveness/Risks/Costs
Examine	Name of Inspection/Purpose of inspection/Inspection method/Inspection Result/ Interpretation/Cost
Medical Technology Department	Section Name/Function/Equipment/Staff
Outpatient Clinical Departments	Section Name/Function/Staff
Prognosis	Prognostic Assessment/Prognostic Factors/Prognostic indicators/Prognostic time/Prognostic results

Table 3: RJUA-QA Dataset Q&A Pairs.

type of task	language	train	val	test
QA	Chinese	1705	211	213

- **Structured Representation:** Extracted knowledge is stored in a structured format with categories. In hypergraph-based KGs, each entity and relationship is a node or hyperedge. This makes it easier to store and use knowledge.

4.2.3 Knowledge fusion. Knowledge Fusion involves the integration of knowledge from multiple sources into a unified and consistent knowledge base. It mainly includes:

- **Data Deduplication:** In multiple data sources, the same entity may have various representations. The goal of data deduplication is to identify and map these duplicate entities and relationships to a single, normalized representation in the KG, ensuring uniqueness and consistency.
- **Conflict Resolution:** When information from different sources is inconsistent, conflict resolution strategies become crucial. Resolving these conflicts usually requires defining a set of rules, such as prioritizing based on the credibility of data sources, using statistical methods for voting, or employing domain experts for arbitration.
- **Knowledge Updating:** Knowledge updates can be performed incrementally, adding or modifying only the parts that have changed since the last update, or through full updates, which involve re-integrating the latest versions of all data sources.

5 Implementation

5.1 Datasets

We use the RJUA-QA dataset, developed by AntGroup Medical LLM and the Department of Urology at Shanghai Jiao Tong University School of Medicine Affiliated Renji Hospital. It has 2,132 questions and answers, as shown in Table 3. Each pair is made up of a question, answer, disease, treatment and reasoning. It covers 97.6% of urological cases. The RJUA-QA data is compiled by doctors and does not involve any personal patient information.

After splitting the specific fields, the specific data of this dataset is shown in Table 4.

5.2 RJUA-HKG construction

Knowledge Extraction: The Context field in the dataset contains unstructured data that is not related to other question-answer information. The extraction process is divided into two stages: QA text data and Context text data.

- **RJUA-QA Text Data:** This includes patient IDs, QA question-answer pairs, disease diagnosis results, and treatment plans.
- **RJUA-Context Text Data:** This consists of inference context information, patient IDs, and Context fields.

In this experiment, entity and relationship extraction is based on the OneKE model, the formal definition of the knowledge extraction task is as follows:

$$K_\theta(Y|RJUA, P, S) = \sum_{i=1}^{n} K_\theta(y_i|RJUA_i, P, S) \quad (1)$$

where $RJUA$ is the extracting context, P denotes prompt, S is the schema, θ represents the model parameters, $Y = (y_1, y_2, \ldots, y_n)$ is the extracted knowledge based on prompt P satisfying schema S, n denotes the number of S relations.

To improve the model's performance in extracting medical knowledge, given the lack of reliable descriptions of medical data in the IEPile dataset, we design an incremental training method as follows:

Table 4: RJUA-QA Dataset Q&A Pairs Character length for each field.

Subset of dataset fields	RJUA-QA				
	Question	Context	Answer	Disease	Advice
Average character length	186	1744	218	8	33
Maximum character length	667	8124	695	68	124

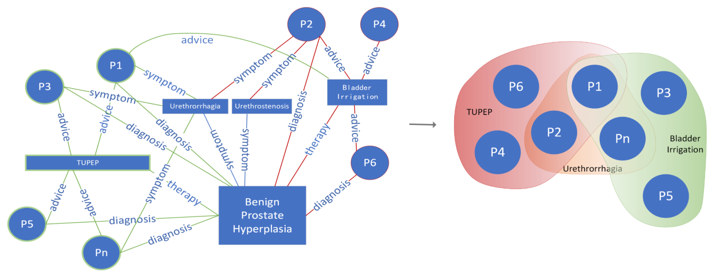

Figure 4: Schematic diagram of the triple KG of urology and urology KG based on a hypergraph.

- **Initial Extraction:** First, we use the OneKE to perform relationship extraction on a portion of the data. The aim is to simultaneously identify entities and their corresponding relationships, generating an initial set of extracted results.
- **Error Correction through Sampling:** Next, we randomly sample from the incorrectly extracted results and manually correct them. These corrected samples form a new dataset used for further model refinement.
- **Fine-tuning with LoRA:** The OneKE is then fine-tuned using the LoRA technique. This process is repeated iteratively until the extraction accuracy across the entire dataset exceeds 95%.

In the construction of the fine-tuning dataset, we introduce contrastive learning by adding a small number of negative samples. This strategy helps increase the distance between positive and negative samples in the dataset, allowing the model to more quickly learn the correct method for extracting knowledge. By reinforcing correct extractions through this contrastive learning approach, we ensure that the model improves its performance on knowledge extraction.

Hyper-edges Aggregation: To integrate the diverse medical knowledge from the RJUA dataset, which includes both patient dialogues and disease-related data, we need to align entities and aggregate relationships by establishing hyperedges. This process involves linking patient complaints and diagnostic treatments through hyperedges based on the doctors' diagnostic results. The goal is to unify and map these to standardized entities, creating a cohesive knowledge representation: $X = \{x_1, x_2, \ldots, x_n\} \longrightarrow ett\_name$, where X encompasses multiple expressions of the same entity. For example, {Benign prostatic hyperplasia, Enlarged prostate, BPH, Benign Prostatic Hyperplasia } -> BPH.

After entity alignment, the diagnostic results of each patient in RJUA-QA are first matched with the disease names described in RJUA-Context, establishing hyperedge connections.

The difference between the graph structure and the hypergraph-based approach is as shown in Figure 4.

Table 5: Results of cleaning the dataset.

Data Set	before cleaning	after cleaning
RJUA-QA	23k	22k
RJUA-Context	9k	1.3k

5.3 Experimental results and discussion

Many triples are found after OneKE model extraction using continuous contrastive learning in RJUA data. The data also needs cleaning because the patient's disease is in the RJUA-Context. The cleaned data is in Table 5.

The experimental results and time complexity are shown in Table 6.

In the KG, we need to compare the symptoms of all patients to find those with similar test indicators. If there are n patients with m symptoms, we need to compare n*m times. Hyper-edge querying only needs to check the serial number of the patient with the most repetitions in the hyper-edge, so it only needs to be m times. When searching for treatment plans for patients with similar diseases, it is also necessary to search for all patients to ensure that the most similar indicators are found. The search in the hyperedge only needs to focus on the number of symptoms, which is usually much smaller than the number of patients.

6 Conclusion

We propose a four-layer knowledge organization method based on hypergraph theory, aiming to address the limitations of existing triple representation methods in handling complex clinical data. By incorporating hypergraph theory and a narrative layer into the traditional knowledge organization framework, this method significantly enhances the representation of higher-order relationships involving multiple entities, particularly in conveying procedural knowledge in medical scenarios. Experiments conduct on a real QA dataset in urology validate the effectiveness of the proposed

Table 6: The results and time complexity.

Scenarios	Number of hypergraph tuples/total tuples	ordinary query	time complexity
Find patients with similar test indicators	3k/23k	ordinary edge query	O(n*m)
		hyperedge query	O(m)
Find treatment options for patients with similar diseases		ordinary edge query	O(n)
		hyperedge query	O(k«n)

RJUA-HKG. The results demonstrate that the model successfully reduces relational complexity, shortens knowledge retrieval time, and accurately captures the dynamic nature of diagnostic and treatment processes.

Future research can further explore the broad applications of this method in the medical field, particularly in downstream applications such as question-answering systems and diagnostic support. Additionally, extending this approach to other medical subfields could be considered to verify its generalizability and adaptability.

Acknowledgments

This research was supported by The National Key Research and Development Program of China under Grants 2022YFC3300801 and The Natural Science Foundation of Hubei Province Educational Committee under Grants B2019009.

References

[1] Dingpu Shi, Jincheng Zhou, Dan Wang, *et al.* Research status, hotspots, and evolutionary trends of intelligent education from the perspective of knowledge graph[J]. Sustainability, 2022, 14(17): 10934.
[2] Samreen Zehra, Syed Farhan Mohsin, Shaukat Wasi, *et al.* Financial knowledge graph based financial report query system[J]. IEEE Access, 2021, 9: 69766-69782.
[3] Jiao Chen, Luyi Ma, Xiaohan Li, *et al.* Knowledge graph completion models are few-shot learners: An empirical study of relation labeling in e-commerce with llms. arXiv preprint arXiv:2305.09858, 2023.
[4] Elizabeth Davidson, Aaron Baird, Karl Prince. Opening the envelope of health care information systems research. Information and Organization[J], 2018, 28(3): 140-151.
[5] Pijush Kanti Dutta Pramanik, Saurabh Pal, Moutan Mukhopadhyay. Healthcare big data: A comprehensive overview[J]. Research anthology on big data analytics, architectures, and applications, 2022, 119-147.
[6] Pin Ni, Ramin Okhrati, Steven Guan, *et al.* Knowledge graph and deep learning-based text-to-GraphQL model for intelligent medical consultation chatbot[J]. Information Systems Frontiers, 2024, 26(1): 137-156.
[7] Nikola Milosevic, Wolfgang Thielemann. Comparison of biomedical relationship extraction methods and models for knowledge graph creation[J]. Journal of Web Semantics, 2023, 75: 100756.
[8] Jean-Paul Calbimonte, Andrei Ciortea, Timotheus Kampik, *et al.* Autonomy in the age of knowledge graphs: vision and challenges[J]. Transactions on Graph Data & Knowledge, 2023, 1(1): 22p.
[9] Xuehong Wu, Junwen Duan, Yi Pan, *et al.* Medical knowledge graph: Data sources, construction, reasoning, and applications[J]. Big Data Mining and Analytics, 2023, 6(2): 201-217.
[10] Yuhuan Lu, Dingqi Yang, Pengyang Wang, *et al.* Schema-Aware Hyper-Relational Knowledge Graph Embeddings for Link Prediction[J]. IEEE Transactions on Knowledge and Data Engineering, 2023.
[11] Jiapu Wang, Boyue Wang, Meikang Qiu, *et al.* A survey on temporal knowledge graph completion: Taxonomy, progress, and prospects[J]. arXiv preprint arXiv: 2308.02457, 2023.
[12] Jianfeng Wen, Jianxin Li, Yongyi Mao, *et al.* On the Representation and Embedding of Knowledge Bases beyond Binary Relations [C]//Proceedings of the 25th International Joint Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2016, 1300-1307.
[13] Bahare Fatemi, Perouz Taslakian, David Vazquez, *et al.* Knowledge hypergraphs: prediction beyond binary relations[C]// Proceedings of the 29th International Joint Conferences on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2021, 2191-2197.
[14] Jonas L.Juul, Austin R.Benson, Jon Kleinberg. Hypergraph patterns and collaboration structure[J]. Frontiers in Physics, 2024, 11: 1301994.
[15] Pei Li, Sheng Wang, Hantao Zhao, *et al.* IG-Net: An interaction graph network model for metro passenger flow forecasting[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(4): 4147-4157.
[16] Bilal Khan, Jia Wu, Jian Yang, *et al.* Heterogeneous hypergraph neural network for social recommendation using attention network[J]. ACM Transactions on Recommender Systems, 2023.
[17] Chaoran Cui, Xiaojie Li, Chunyun Zhang, *et al.* Temporal-relational hypergraph tri-attention networks for stock trend prediction[J]. Pattern Recognition, 2023, 143: 109759.
[18] Feng Gao, Lili Zheng, Jinguang Gu. Financial Knowledge Graph Representation and Construction for Multiple Time-Series Relationships[J]. Journal of Shanxi University (Natural Science Edition), 2022.
[19] Zhe Chen, Yuehan Wang, Bin Zhao, *et al.* Knowledge graph completion: A review[J]. IEEE Access, 2020, 8: 192435-192456.
[20] Basra Jehangir, Saravanan Radhakrishnan, Rahul Agarwal. A survey on Named Entity Recognition—datasets, tools, and methodologies[J]. Natural Language Processing Journal, 2023, 3: 100017.
[21] Diaobo Sui, Yubo Chen, Kang Liu, *et al.* Joint entity and relation extraction with set prediction networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023.
[22] Tianshuo Peng, Zuchao Li, Lifei Zhang, *et al.* FSUIE: A novel fuzzy span mechanism for universal information extraction[J]. arXiv preprint arXiv:2306.14913, 2023.
[23] Honghao Gui, Shuofei Qiao, Jintian Zhang, *et al.* InstructIE: A Bilingual Instruction-based Information Extraction Dataset[J]. arXiv preprint arXiv:2305. 11527, 2023.
[24] Hugo Touvron, Louis Martin, Kevin Stone, *et al.* Llama 2: Open foundation and fine-tuned chat models[J]. arXiv preprint arXiv:2307.09288, 2023.
[25] Alain Bretto. Hypergraph theory[J]. An introduction. Mathematical Engineering. Cham: Springer, 2013, 1.