Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 9
AIDML 2025

Knowledge Graphs: Technical Construction, Cross-Domain Applications and Future Challenges

Xuanyu Chen*

The College of Mechanical Engineering, Donghua University, Shanghai, China

*Corresponding author: 230860108@mail.dhu.edu.cn

Abstract. The Knowledge Graph has become a crucial technology in the field of Artificial Intelligence, structurally representing real-world entities and their semantic relationships through the Semantic Web. The utilization of knowledge graphs has been demonstrated to facilitate the enhancement of semantic understanding and knowledge extraction for computers. Despite the wide applicability of knowledge bases, there are still challenges in dealing with complex knowledge types and the accuracy of large-scale knowledge bases. This paper systematically reviews the fundamental concepts, construction methodologies, and representative application scenarios of knowledge graphs, including recommender systems, question answering systems, and educational systems. Furthermore, it examines how knowledge graphs are developing right now and considers possible future study avenues based on previous findings. By synthesizing theoretical foundations with practical perspectives, this paper aims to facilitate deeper understanding, effective implementation, and continuous innovation in knowledge graph technologies across a wide range of domains. This research can provide both theoretical foundations and practical insights for researchers and practitioners seeking to understand, implement, or advance knowledge graph technologies across various domains.

Keywords: Knowledge graph; Knowledge graph construction; graph embedding; Recommender systems; Question answering systems.

1. Introduction

With the rise of Large Language Models (LLMs) such as ChatGPT [1], artificial intelligence has attracted widespread attention. Knowledge graph (KG) is an essential technology in the field of artificial intelligence. The knowledge graph can be thought of as the development and improvement of the semantic web in the age of big data on the internet [2]. Research on important technologies like knowledge representation, knowledge acquisition, knowledge fusion, and knowledge reasoning has been a focus of industry and academics in the development of knowledge graphs. The primary function of knowledge graph organization creation and question and answer retrieval is knowledge extraction, which is crucial for deep semantic processing and understanding. By creating links between entity concepts, knowledge graphs assist computers in comprehending the nature of data and, consequently, in explaining the characteristics of real-world objects and phenomena.

Many different fields have made extensive use of knowledge graphs, such as semantic search, machine question and answer, and online learning [3]. By utilizing the entity relationships in the knowledge graph, machines can more accurately understand and answer users' questions. Meanwhile, the application of knowledge graphs in social networks and intelligent recommendation systems is also emerging, providing users with more personalized and intelligent experiences [4]. Furthermore, as technology develops, knowledge graphs are being combined more and more with deep learning methods to strengthen their capacity for reasoning and increase the precision of knowledge extraction in order to investigate novel applications in domains such as healthcare [5]. Nevertheless, there is still work to be done to address the challenge of knowledge graphs when dealing with complex knowledge types and large knowledge bases that contain error triples [6].

In this paper, the notion of a knowledge graph is presented, including categories and construction of knowledge graph, application domains with examples of recommender systems, intelligent Q&A

systems and educational systems, as well as current limitations and future perspectives of knowledge graphs, aiming to help readers quickly understand knowledge graph or provide useful references for the researchers in the related fields.

2. Basic Theory

2.1. Definitions and Categories

Google first proposed the idea of knowledge graph in 2012 [6]. The knowledge graph is a structured database for recording relationships between various items [7]. Real-world elements and their interactions are represented in a knowledge graph as directed graphs. Knowledge graphs are based on the idea that artificial intelligence symbols can express a lot of information in the real world within the framework of the human mind. This suggests that they can explain a wide range of attributes, concepts, and relationships that may exist in the actual world and combine them to create a vast semantic learning network graph, where each node represents an entity or an idea [8]. It is composed of entities, relations and attributes, represented as a triad (head entity, relation, tail entity). In the knowledge graph, Nodes represent entities of interest, while edges reflect interactions between entities. The edge between two entities represents their semantic relationship. The direction of a link is essential as the relationship isn't always symmetrical [9]. According to their states, static and dynamic knowledge graphs are two types of knowledge graphs. The data of a static knowledge graph is relatively fixed, reflecting the state of knowledge up to a certain moment or a specific time period, which is mainly used to organize and represent known information. In contrast, dynamic knowledge graphs are updated in real time. The most typical is the temporal knowledge graph. It adds a temporal dimension to the traditional knowledge graph to capture the evolution and development of relationships over time. Because of its near real-time updating characteristics, dynamic knowledge graphs are mostly used to deal with rapidly changing domains. Knowledge graphs can be divided into two categories based on their domain: generic knowledge graphs and specific domain knowledge graphs. The general knowledge graph contains complex general knowledge in various domains, focusing on the breadth of knowledge. Conversely, a specific domain knowledge graph focuses on the depth of knowledge, requiring a high degree of expertise and accuracy [10].

2.2. Knowledge Graph Construction

The construction of a knowledge graph is predicated on the acquisition of data, encompassing structured, semi-structured, and unstructured data. There are two main methods for constructing knowledge graphs: bottom-up and top-down. After the data has been collected, utilize models like Bert to identify the entities and their relationships inside the fragmented text [11]. Due to the diversity and heterogeneity of data, there is a need for knowledge fusion so that the knowledge graph integrates a unified form [12]. In order to circumvent the issue of coincident entities across disparate data sources, ontology alignment is employed to identify and map similar concepts. Subsequently, the identification and resolution of redundant information and data conflicts is necessary, followed by their integration into a unified knowledge graph. To make it easier for computers to process and represent knowledge, a variety of models are employed to transform information into a format that is processable by computers [13]. A few examples of frequently used representations are those based on graphs, triples, distance models, semantic models, etc [14]. Finally, the accuracy and reliability of the triples can be evaluated through the external knowledge base [15].

3. Areas of Application

3.1. Recommender System

References are cited in the text just by square brackets [1]. (If square brackets are not available, slashes may be used instead, e.g. /2/.) Two or more references at a time may be put in one set of brackets [3, 4]. The references are to be numbered in the order in which they are cited in the text and are to be listed at the end of the contribution under a heading References, see the example below.

$$\forall U_i \in U, v_k = \underset{v_j \in V}{\operatorname{argmax}} f\left(u_i, v_j\right) \tag{1}$$

Mainstream recommendation algorithms can be categorized into content-based recommendation, collaborative filtering (CF)-based recommendation and hybrid recommendation systems [16]. The principle of content-based recommender systems is to analyze the user's feature preferences (e.g., tags, categories) and recommend items with similar characteristics. Predicting a user's interest based on past behavior (such as ratings or purchases) or item correlation is the idea behind CF-based recommender systems. Both item-based collaborative filtering and user-based collaborative filtering fall under this category [17]. In order to address the cold start and sparsity problems in the two recommendation systems mentioned above, hybrid recommender systems have been developed. Hybrid recommendation models can make up for an algorithm's shortcomings and strengthen the advantages of other algorithms. The CF algorithm in combination with other algorithms is the most widely used.

Conventional recommender systems are only capable of providing the outcomes of recommended items, which may result in diminished user recognition. The introduction of KG has been demonstrated to engender an enhancement in the interpretability of recommended items. Knowledge graphs (KGs) store rich semantic relationships, from which recommender systems can derive connections between items in order to facilitate a more profound comprehension of users' interests and preferences, providing more personalized recommendations. They fall into three primary categories: approaches based on embeddings, approaches based on paths, and approaches based on propagation [18]. Two-stage learning methods, joint learning methods, and alternate learning methods are the three further groups into which embedding-based approaches can be separated [19]. Entities and relationships are mapped to a low-dimensional vector space using knowledge graph embedding (KGE) [20]. TransE is the most basic model of KGE. The fundamental principle asserts that the head node h, the relation r, and the tail node t of a triple possess corresponding vectors. The equation $h + r = t$ can be derived from the triple. Such an equation is designated a fact. TransE expects $h + r \approx t$, believing the closer the result of $h + r$ is to t, the more plausible the recommendation is, and the more likely the entities and relationships in KG are to be established [21]. Derived from this are algorithms such as TransH [22], TransR [23], and hybrid algorithms such as TransE-MTP [24]. Path-based approaches make recommendations by analyzing paths between users and items in the knowledge graph and learning connection similarities [25]. PathSim is frequently used to quantify how similar the relationships between entities in a graph are [26]. It is described as

$$s_{x,y} = \frac{2 \times \left|\{p_{(x,y)}: p_{(x,y)} \in \mathcal{P}\}\right|}{\left|\{p_{(x,x)}: p_{(x,x)} \in \mathcal{P}\}\right| + \left|p_{(y,y)}: p_{(y,y)} \in \mathcal{P}\right|} \tag{2}$$

where $p_{(x,y)}$ is a path from entity x to entity y. To enhance the recommendation outcomes, the user-user, entity-entity, and user-entity similarity are all regularized by the Hete-CF algorithm [27]. Propagation-based approaches mainly utilize the relationships between entities in KG to propagate user interests or item features to related entities. Graph Neural Networks (GNN) serve as the foundation for common implementations. Graph neural networks' neighbor aggregation approach can capture multi-hop semantic propagation in knowledge graphs [28]. The typical model is Knowledge Graph Attention Network (KGAT), where the embedding update of each node depends on the weighting information of its neighbors [29].

3.2. Question Answering(QA) System

Humanity's thirst for knowledge drives the development and progress of society. In order to make it easier and more accurate to get the information, QA systems have been created. By comprehending user queries and extracting or deriving precise responses from vast volumes of data, QA systems are made to deliver effective and practical information services. With large amounts of data and rich semantic information, KG can improve the accuracy and efficiency of QA systems to answer more complex questions [30]. Knowledge graph-based question answering (KGQA) can be categorized into methods that rely on information retrieval (IR) and semantic parsing (SP) [31]. SP-based methods convert natural language questions into logical forms. These logical forms are then transformed into knowledge graph query statements. Natural semantic understanding and graphical databases can be used to make the Q&A process simpler and more efficient [32]. IR-based method turns the questions into a text matching task in which candidate paths are generated and ranked to find answers [33]. The entity recognition phase employs Bidirectional LSTM (BiLSTM) coupled with Conditional Random Fields (CRF) to identify the topic entity, followed by entity linking for precise knowledge graph anchoring. Subsequently, path templates are leveraged to perform multi-hop relational inference, generating candidate paths. Through a path ranking mechanism that evaluates semantic relevance and structural coherence, optimal paths are selected and fused via knowledge-aware aggregation to derive the final answer. Furthermore, graph-metric-based approach can be implemented by embedding triples (s, r, o) into a continuous low-dimensional vector space, where entities and relations are represented as e_s, e_r and e_o, respectively. A scoring function $f(e_s, e_r, e_o)$ is then designed to evaluate the plausibility of triples, with the optimization objective of either minimizing the loss function or maximizing optimization objectives [34]. A model called TransKGQA leverages the Transformer architecture to encode contextual information through multi-head self-attention mechanisms [35]. It aligns encoded representations with KG via cosine similarity-based matching, followed by structured Cypher query generation to retrieve precise answers from the graph database.

3.3. Education System

In accordance with technological advances, educational methodologies are subject to constant refinement and enhancement. Knowledge graphs are also being used in education systems to optimize the representation of knowledge and to personalize education. Disciplinary knowledge graph is one of the examples. Knowledge points (such as formulas and concepts) and related learning resources (such as exercises and course materials) make up the nodes of a discipline knowledge graph. Discipline knowledge graphs can create learning repositories that are customized for each student by utilizing natural language processing techniques. They can model learner profiles based on individual learner characteristics, while dynamically perceiving and updating external knowledge [36]. When students' prior knowledge, learning styles, and interests are known, by analyzing the logic of what students already know, personalized learning path recommendation system can suggest appropriate learning routes for them [37]. The relationship between descriptive language and instructional resource knowledge points was determined using a semantic similarity algorithm.

$$\text{Match}(R, K) = S(D_R, K) \tag{3}$$

R stands for a particular teaching resource, K stands for a specific knowledge point, and D_R is the R description of the resource. Corresponding knowledge points are automatically linked to highly relevant sites. The system, named KnowEdu, is capable of constructing a knowledge graph by utilizing heterogeneous data from the field of education. It employs a neural sequence labeling algorithm to extract teaching concepts and applies a probabilistic association rule mining method to extract data for learning assessment [38].

4. Conclusion

Knowledge graphs serve as a cornerstone for semantic understanding and reasoning across diverse domains, enabling structured representation of entities and their relationships. Their applications span several areas, first is recommender systems, where KGs enhance interpretability by modeling user-item interactions through semantic linkages, addressing cold-start and data sparsity issues through hybrid approaches; Second is question answering (QA) systems, which leverage KGs to improve answer accuracy via multi-hop relational inference and graph-based query parsing, allowing systems to handle complex, context-dependent queries; and Third is educational systems, where KGs facilitate personalized learning path recommendations by mapping knowledge dependencies and learner profiles, dynamically adapting to individual learning styles and interests. Advanced techniques such as KGE and graph neural networks (GNNs) further empower AI-driven tasks by encoding latent semantics and propagating contextual information across graph structures, enabling richer representations and more accurate predictions.

Despite these advancements, critical challenges persist, including scalability limitations in managing large-scale, error-prone knowledge bases, and the integration of dynamic, heterogeneous data sources. Future research should prioritize synergizing KGs with emerging technologies like LLMs and multimodal frameworks to improve reasoning capabilities and ensure real-time knowledge updates. Additionally, domain-specific applications requiring robust foundational knowledge bases—such as healthcare, finance, and smart cities—present untapped opportunities to expand KG utility. Addressing these challenges will not only strengthen the foundational role of KGs in AI but also unlock their potential in solving complex real-world problems.

References

[1] Mohammad A. F., Clark B., Hegde R. Large Language Model (LLM) and GPT, a Monolithic Study in Generative AI. In: 2023 Congress in Computer Science, Computer Engineering, and Applied Computing (CSCE), IEEE, 2023: 383-388.

[2] T. A Monolithic Study in Generative AI. 2023 Congress in Computer Science, Computer Engineering, and Applied Computing (CSCE), Las Vegas, USA, 2023: 383-388.

[3] Yahya M., Breslin J. G., Ali M. I. Semantic Web and Knowledge Graphs for Industry 4.0. Applied Sciences, 2021, 11(11): 5110.

[4] Yu Mengbo, Du Jianqiang, Luo Jigen, Nie Bin, Liu Yong, Qiu Junyang. Research Progress of Knowledge Graph Completion Based on Knowledge Representation Learning[J]. Computer Engineering and Applications, 2023, 59(18): 59-73.

[5] Yu Z., Xu Y. A Review of Research on the Application of Knowledge Graphs in Recommender Systems. 2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2024: 1843-1849. doi: 10.1109/ITNEC60942.2024.10733130.

[6] Zheng Z., Liu Y., Zhang Y., Wen C. TCMKG: A Deep Learning Based Traditional Chinese Medicine Knowledge Graph Platform. 2020 IEEE International Conference on Knowledge Graph (ICKG), Nanjing, China, 2020: 560-564. doi: 10.1109/ICBK50248.2020.00084.

[7] Yang Y., Chen J., Xiang Y. A Review on the Reliability of Knowledge Graph: From a Knowledge Representation Learning Perspective. World Wide Web, 2025, 28(4).

[8] Badrouni M., Katar C., Inoubli W. Large-Scale Knowledge Graph Representation Learning. Knowledge and Information Systems, 2024, 66: 5479–5499.

[9] Wu Y., Fu X. F., Xu L., Jiang Z. F. A Summary of the Latest Research on Knowledge Graph Technology. 2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), Shanghai, China, 2021: 38-42. doi: 10.1109/ICCEAI52939.2021.00007.

[10] Lin J., Zhao Y., Huang W. Domain Knowledge Graph-Based Research Progress of Knowledge Representation. Neural Computing and Applications, 2021, 33: 681–690.

[11] Peng C., Xia F., Naseriparsa M., etc. Knowledge Graphs: Opportunities and Challenges. Artificial Intelligence Review, 2023, 56: 13071–13102.

[12] Huang Y., Yu S., Chu J., Su Z., Cong Y. Combining Deep Learning with Knowledge Graph for Design Knowledge Acquisition in Conceptual Product Design. Computer Modeling in Engineering and Sciences, 2024, 138(1): 167–200.

[13] Zhao X., etc. Multi-Source Knowledge Fusion: A Survey. 2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC), Hangzhou, China, 2019: 119-127.

[14] Lin J., Zhao Y., Huang W. Domain Knowledge Graph-Based Research Progress of Knowledge Representation. Neural Computing and Applications, 2021, 33: 681–690.

[15] Li F., Xie W., Wang X., Fan Z. Research on Optimization of Knowledge Graph Construction Flow Chart. 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 2020: 1386-1390.

[16] Ban T., Wang X., Chen L., Wu X., Chen Q., Chen H. Quality Evaluation of Triples in Knowledge Graph by Incorporating Internal With External Consistency. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(2): 1980-1992.

[17] Adomavicius G., Tuzhilin A. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6): 734-749.

[18] Alharthi H., Inkpen D. Content-Based Recommender System Enriched with Wordnet Synsets. In: Gelbukh A. (eds) Computational Linguistics and Intelligent Text Processing. CICLing 2015. Lecture Notes in Computer Science, vol 9042. Springer, Cham.

[19] Guo Q., etc. A Survey on Knowledge Graph-Based Recommender Systems. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(8): 3549-3568.

[20] Zhang Jin-Cheng, Zain Azlan Mohd, Zhou Kai-Qing, Chen Xi, Zhang Ren-Min. A Review of Recommender Systems Based on Knowledge Graph Embedding. Expert Systems with Applications, 2024, 250: 123876.

[21] Wang Hui-xin, Tong Xiang-rong, Research Progress of Recommendation System Based on Knowledge Graph, Journal of Zhejiang University (Engineering Science), 57(8), 2023, pp. 1527–1540.

[22] Bordes A., Usunier N., Garcia-Duran A., Translating Embeddings for Modeling Multi-Relational Data, Advances in Neural Information Processing Systems, 26, 2013, pp. 2787–2795.

[23] Zhen W., Zhang J., Feng J., Knowledge Graph Embedding by Translating on Hyperplanes, Proceedings of the National Conference on Artificial Intelligence, Québec, Canada, AAAI Press, 2014, 28(1), pp. 1112–1119.

[24] Li Y., Zhu C., TransE-MTP: A New Representation Learning Method for Knowledge Graph Embedding with Multi-Translation Principles and TransE, Electronics, 13(16), 2024, Article 3171.

[25] Luo Chengtian, Ye Xia, Survey on Knowledge Graph-Based Recommendation Methods, Computer Engineering and Applications, 59(1), 2023, pp. 49–60.

[26] Sun Y., Han J., Yan X., Yu P. S., Wu T., PathSim: Meta Path-Based Top-K Similarity Search in Heterogeneous Information Networks, Proceedings of the VLDB Endowment, 4(11), 2011, pp. 992–1003.

[27] Yu X., Ren X., Gu Q., Sun Y., Han J., Collaborative Filtering with Entity Similarity Regularization in Heterogeneous Information Networks, Proceedings of the 1st International Joint Conference on Artificial Intelligence Workshop on Heterogeneous Information Network Analysis, 2013.

[28] Wu Shiwen, Sun Fei, Zhang Wentao, Xie Xu, Cui Bin, Graph Neural Networks in Recommender Systems: A Survey, ACM Computing Surveys, 55(5), Article 97, 2023, 37 pages.

[29] Wang Q., Cui H., Zhang J., Neighbor-Augmented Knowledge Graph Attention Network for Recommendation, Neural Processing Letters, 55, 2023, pp. 8237–8253.

[30] Wu L., Yang L., Zhang C., Xu Y., Research on Application and Optimization Model of Knowledge Graph in Intelligent Question Answering System, 2024 IEEE 2nd International Conference on Electrical, Automation and Computer Engineering (ICEACE), Changchun, China, 2024, pp. 1686–1690, doi: 10.1109/ICEACE63551.2024.10898299.

[31] Wu P., Zhang X., Feng Z., A Survey of Question Answering over Knowledge Base, Knowledge Graph and Semantic Computing: Knowledge Computing and Language Understanding, 4th China Conference, CCKS 2019, Hangzhou, China, 2019, Revised Selected Papers, Springer, pp. 86–97.

[32] Varma S., Shivam S., Biswas S., Saha P., Jalan K., Graph NLU Enabled Question Answering System, Heliyon, 7(9), 2021, Article e08035.

[33] Xiao Y., Zhang L., Huang J., Zhang L., Wan J., An Information Retrieval-Based Joint System for Complex Chinese Knowledge Graph Question Answering, Electronics, 11(19), 2022, Article 3214.

[34] Song Y., Li W., Dai G., Shang X., Advancements in Complex Knowledge Graph Question Answering: A Survey, Electronics, 12(21), 2023, Article 4395.

[35] Li Chong Y., Poo Lee C., Zen Muhd-Yassin S., Ming Lim K., Kamsani Samingan A., TransKGQA: Enhanced Knowledge Graph Question Answering with Sentence Transformers, IEEE Access, 12, 2024, pp. 74872–74887, doi: 10.1109/ACCESS.2024.3405583.

[36] Zhao Yubo, Zhang Liping, Yan Sheng, Hou Min, Gao Mao, Construction and Application of Discipline Knowledge Graph in Personalized Learning, Computer Engineering and Applications, 59(10), 2023, pp. 1–21.

[37] Li H., Duan W., Construction and Application Research of Intelligent Education Knowledge Graph Based on Multi-Modal Learning, 2024 International Conference on Informatics Education and Computer Technology Applications (IECA), Beijing, China, 2024, pp. 117–121, doi: 10.1109/IECA62822.2024.00029.

[38] Chen P., Lu Y., W V., Zheng, Chen X., Yang B., KnowEdu: A System to Construct Knowledge Graph for Education, IEEE Access, 6, 2018, pp. 31553–31563.