IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 7 February 2024, accepted 22 February 2024, date of publication 26 February 2024, date of current version 5 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3370409

TOPICAL REVIEW

A Bibliometric Analysis of Recent Developments and Trends in Knowledge Graph Research (2013–2022)

GANG WANG AND JING HE
School of Foreign Languages, Chaohu University, Chaohu 238000, China
Corresponding author: Gang Wang (057052@chu.edu.cn)

This work was supported in part by the Humanity and Social Science Project of Anhui Province under Grant SK2019A0541, and in part by the Project of Quality Engineering of Anhui Province under Grant 2020kfkc339.

ABSTRACT Knowledge graphs have emerged as a useful resource and tool for representing real-world entities and their relations, which gained increasing importance in the fields of deep learning and machine learning. This research aims to investigate the academic publications of knowledge graphs between 2013 and 2022 based on the core collection of the Web of Science and examine hot topics and the latest developments in this subject. Thus, the present research adopted a bibliometric analysis to explore the indicators, which mainly focus on different variables from the diachronic productivity of scientific publications to the most prolific countries and the leading publication journals. By means of VOSviewer software, the most productive authors and the frequency of author keywords were further analyzed. The results manifest that dramatic growth has been identified in the past five years due to the output of publications regarding this subject, and the frequently explored themes were mainly conducted from six dimensions, focusing on ontology modelling, knowledge extraction, knowledge graph embedding, graph-based knowledge representation, multi-modal knowledge graphs and knowledge-aware applications. The findings could help researchers gain a thorough understanding of knowledge graph research, optimize research topic choices, and detect new directions for future studies.

INDEX TERMS Bibliometric analysis, knowledge graphs, VOSviewer.

I. INTRODUCTION

Advances in artificial intelligence (AI) are transforming the ways people acquire knowledge, and the effective application of semantic technology profoundly influences traditional knowledge representation modes. Knowledge is often viewed as information with specific properties, which has been utilized to deal with the extant human knowledge structures in an integrated way [1]. The approaches to identifying and representing useful knowledge from abundant information have garnered significant attention from academia and industry. Together with deep learning, knowledge graphs unfold various techniques that could be adopted to integrate and extract effective information based on multiple data sources. With great competence in knowledge representation, knowledge graphs have found a wide utilization in a variety of fields, demonstrating their advantages in vast task-specific applications [2], such as semantic retrieval, language representation learning and question answering systems.

A knowledge graph, a structured knowledge representation via graphs, could date back to expert systems developed in the 1970s [3]. In the 1980s, researchers initially introduced knowledge graphs to describe their knowledge-based systems [4]. In the present study, the term knowledge graph originates from the emergence of Google Knowledge Graph, and since the year 2012, the phenomenon of knowledge graphs has become known worldwide as it enables users to look for objects, persons, or locations [5]. Moreover, with the advent of linked data, considerable research into knowledge graphs has been carried out to explore their interpretations and applications.

Currently, although there have been no well-established definitions of knowledge graphs, a basic common understanding of the term can be sorted out. Knowledge graphs

The associate editor coordinating the review of this manuscript and approving it for publication was Pasquale De Meo.

can be considered as a graph of data with the intention of accumulating and transferring knowledge of the real world, and they are shown as entities and relations, which mainly store structured knowledge and unstructured knowledge, and replicate the unique network of information flow in an organization [5]. To be specific, knowledge graphs describe the domain knowledge about entities, relations and attributes. As for a graph-based knowledge representation, each fact contains a collection of triples of the form (h, r, t), and head entity h and tail entity t are in a connection with a specific relation called r (e.g. Socrates, nationality, Greece). Under many circumstances, the collection of relations could be presented as an ontology, which clarifies the correlations or restrictions of their usage.

Given their significance in diverse domains, publications related to knowledge graphs have led to heated discussions in both theoretical discussion and practical implementation. A number of researchers have sought to address the variables from both macro-level reviews of knowledge graphs and micro-level application analyses of specific aspects of knowledge graphs. For instance, Chen et al. [6] conducted a systematic review of knowledge graph completion, which analyzed the existing mainstream methods and tackled the main bottlenecks encountered by knowledge graph completion tasks. In a similar way, Ji et al. [7] presented a practical review of the representation and implications of knowledge graphs, which provided a development direction of overall research topics in knowledge graphs and proposed a full-view categorization of these subjects. Rizun [5] made a literature review of knowledge graph application in education, and emphasized knowledge graphs can be identified as a technology used to facilitate knowledge management and offer a systematic analysis of the knowledge graph didactic process. At the micro-level, typical applications of knowledge graphs have been discussed in different settings, in particular the educational setting. Cui and Yu [8] examined the effects of knowledge graphs on fostering deep learning in a flipped classroom, which indicated that learning with the help of knowledge graphs can achieve better performance in the students' learning products. Similarly, Wu and Jia [9] investigated the concrete construction and the corresponding applications of the English major-specific knowledge graphs through extensive data sources in an educational domain and verified the interconnection of nodes within structured knowledge graphs. In summary, to provide a more intuitive presentation of the knowledge graph review studies, a list of existing review articles is presented in Table 1.

In general, previous studies on knowledge graphs commonly employed qualitative research methodologies with relatively limited data [6], [7], [15]. Most of them expounded on the architectures and key techniques of knowledge graphs. As far as we know, attempts with a bibliometric approach to examine the scientific publications of knowledge graphs are still scarce, so it is of significance to conduct a diachronic survey of knowledge graphs via a bibliometric study. Therefore, to implement this study, we attempt to

TABLE 1. Summary of knowledge graph review articles.

Year	Authors	Title
2017	Wang et al. [10]	Knowledge graph embedding: a survey of approaches and applications
2018	Yan et al. [11]	A retrospective of knowledge graphs
2020	Chen et al. [6]	Knowledge graph completion: a review
2021	Issa et al. [12]	Knowledge graph completeness: a systematic literature review
2021	Tiwari et al. [13]	Recent trends in knowledge graphs: theory and practice
2021	Abu-Salih. [14]	Domain-specific knowledge graphs: a survey
2022	Ji et al.[7]	A survey on knowledge graphs: representation, acquisition, and applications

evaluate the recent advances in knowledge graph research between 2013 and 2022, aiming to discover the growing trends of knowledge graph research over the past decade, and reveal the scientific landscape of knowledge graphs by the aid of performance analysis and science mapping. Therefore, the four research questions in the present study are listed below.

1. What is the distribution trend of the scientific productions over the examined period in the domain of knowledge graphs?
2. What are the most impactful factors in the domain of knowledge graphs, including countries, funding sponsors, publication journals and authors?
3. What are the most frequently used keywords in the domain of knowledge graphs?
4. What are the current hot topics in the domain of knowledge graphs?

II. METHODOLOGY

A. DATA COLLECTION

Bibliometrics is a quantitative approach employed to tackle scientific data, track scientific advancement, discern research impact, and detect emerging trends [16]. In the field of new technologies, as many review studies were carried out with this approach [17], [18], it has been proved as an objective analysis for measuring the significance of research articles to the deepening of knowledge [19]. Therefore, the present study adopted a bibliometric review and illustrated the procedures of data collection and processing in detail.

To address the research questions, the retrieved information of publications was obtained from the Web of Science core collection, which is known as a reputable and high-cited scientific journals with complete citation records and enhanced metadata [20]. Specifically, the Science Citation Index Expanded, the Social Sciences Citation Index as well as the Arts and Humanities Citation Index were extracted from the authoritative database, and the bibliometric data of all publications needed were downloaded, including article titles, abstracts, author names and keywords.

To bring more validity to the data resources, a meticulous procedure for the data selection was carried out. First of all, the focus of the search string was set up to limit the strings to a few key terms. In view of the possibility of synonyms used in search strings, the search query took similar terms and spelling patterns into consideration, and the combined query was generated as below: TS= ("knowledge graph*" OR "knowledge visualization*" OR "knowledge map*"). Altogether, we tested three related terms on March 16, 2023. Then, the bibliometric information of publications published between January 1, 2013 and December 31, 2022 was considered as the current research sought to observe the trend of development with regard to knowledge graphs. What's more, the document type of the studied publications was restricted to "article". As shown in Table 2, a set of inclusion and exclusion norms was defined to remove the unrelated data [21]. Eventually, after the double check made by the authors, a total of 3058 final samples with their bibliometric information obtained from the WoS were derived in the form of TXT.

TABLE 2. Norms of inclusion and exclusion.

Norms of inclusion	Norms of exclusion
Research is relevant to knowledge graphs	Research is not relevant to knowledge graphs.
The document type is restricted to research articles.	Review article, editorial material and note.
Articles published between 2013 and 2022.	Articles published outside the range between 2013 and 2022.
Only English articles are considered.	Articles written in other languages are not considered.
The full document can be accessible via the subscription.	The full document can not be accessible.

B. DATA ANALYSIS

The retrieved data was further analyzed and categorized by using a visualized software tool. We employed two research instruments to carry out data analysis and their presentation, consisting of ECharts (version 5.1.2) and VOSviewer (version 1.6.18). As an open-source visualization tool, ECharts was utilized to handle the bibliographic information and produce relevant charts. In addition, being a Java-based application, VOSviewer specializes in visual presentation and trend detection of the studied articles [22]. Thus, it was employed to visualize the co-authorship relations and the network of author keywords and in turn produce the bibliometric profiles. Lastly, both text information and bibliographic data were processed and presented in the corresponding charts [20].

III. RESULTS

Answers to the outlined research questions in this section are reported via the data statistics and analysis of the retrieved articles, and they are consistent with the questions mentioned above.

A. DIACHRONIC PRODUCTIVITY OF PUBLICATION

The annual distribution of retrieved articles with regard to knowledge graphs is obtained in Fig. 1. The results indicate the dynamic changes in the number of academic articles published within the examined span. The publication outputs on this subject slightly increased between 2013 and 2017. Beginning with the year 2018 that witnessed 116 articles, the annual productivity was in a growth trend. Moreover, it should be pointed out that the year of 2022 occupies 40.87% of the total production with more than 1200 articles, which may show an increasing interest among researchers. In other words, the research of knowledge graphs has gradually been a heated topic since 2018, which was largely in line with the rapid advancements in the areas of deep learning, machine learning as well as natural language processing. In the short term, the upward trend is more likely to be maintained in the near future based on the increasing outputs of annual publications.

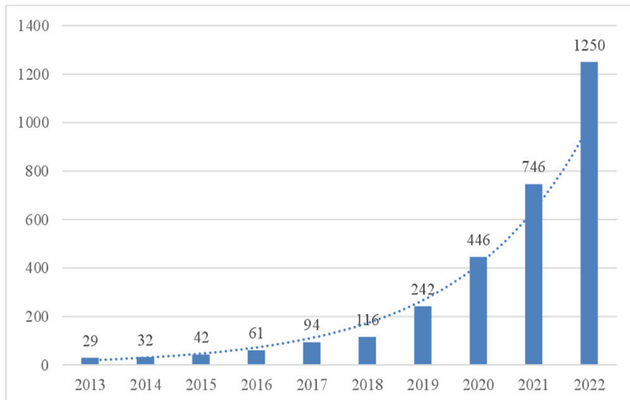

FIGURE 1. Annual publications of knowledge graphs (2013-2022).

B. COUNTRIES DISTRIBUTION AND FUNDING SPONSORS

The results indicate a total of 69 countries have achieved a lot in the area of knowledge graph. According to the first author of the retrieved articles, Table 3 presents the top 10 productive countries in the last decade. In general, Asian countries ranked at the top with the large number of publications (63.17%), while European (21.64%), North American (16.80%), and Oceania (4.12%) countries followed closely behind. As for productive countries, China made fruitful achievements with 1849 articles, followed by the United States (426 articles), the United Kingdom (175), Germany (172) as well as Australia (126). Comparatively speaking, China remained ahead with a high output of publications in knowledge graph research, which takes up more than 60% of all outputs within the research span [23].

Taking funding sponsors into consideration, Fig. 2 presents the major 10 funding sponsors. These sponsors from China as well as America were seen as the two primary sources, among which Chinese agencies have most contributed to knowledge graph research, in particular the National Natural Science Foundation of China (number=1043), the National Key Research and Development Program of China (292). To some extent, a large number of funding programs from

TABLE 3. Top 10 productive countries (2013-2022).

Rank	Countries	Publications	Percentage
1	China	1849	60.46
2	United States	426	13.93
3	United Kingdom	175	5.72
4	Germany	172	5.62
5	Australia	126	4.12
6	Italy	116	3.79
7	Spain	103	3.36
8	France	96	3.13
9	Canada	88	2.87
10	South Korea	83	2.71

China provided strong support for a high output of academic articles in the country. To our knowledge, as the leading agency for the management of China's National Science Fund, the National Natural Science Foundation of China (NSFC) has been dedicated to furthering the advancements of basic research and the construction of fundamental disciplines. So it can be deduced that financial support from governments or agencies seems to be a driving force to boost research productivity in scientific publications [24].

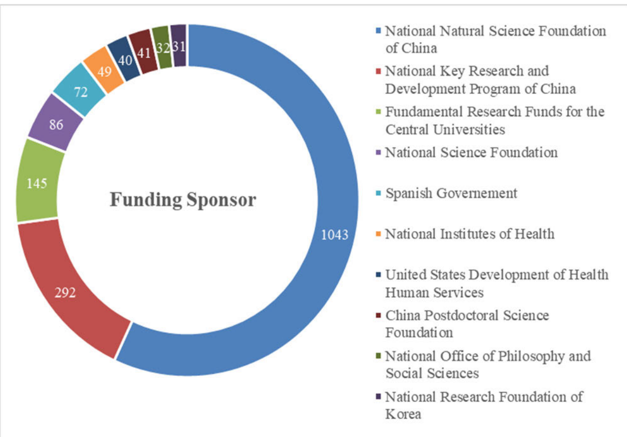

FIGURE 2. Top 10 funding sponsors (2013-2022).

C. PUBLICATION JOURNALS AND SUBJECT CATEGORIES

Based on the publication outputs and quartile rankings in the Journal Citation Report (JCR) for 2022, Table 4 presents the major 10 publication journals with their percentage and corresponding impact factors. These journals published 752 articles, which is up to 24.59% of all publications in this area. Among them, *IEEE Access* (179 articles), *Knowledge-Based Systems* (122) and *Applied Sciences-Basel*(92) are in the top three due to higher publication productivity. It was notable that four of the listed journals ranked in the first quartile by the JCR. Thus, it could be inferred that the publication journals regarding the fields of engineering and AI have paid close attention to knowledge graphs.

Moreover, as the studied articles were distributed in a variety of subjects, Fig. 3 illustrates the top 10 relevant subjects based on the WoS categories. Specifically, computer science and information systems (899 articles, 29.39%) was the most relevant subject, while computer science and artificial intelligence (819, 26.78%), engineering electrical and electronic (518, 16.93%) followed behind. Among these categories, environmental sciences (140, 4.57%) is the least subject category of publication [25]. By considering the related indicators, publications related to the categories of computer science and engineering are considered as the most prominent providers, indicating their strong relevance to knowledge graph research.

TABLE 4. Top 10 publication journals.

Publication journals	Publications	Percentage	Impact factor
IEEE Access	179	5.853	3.9
Knowledge-Based Systems	122	3.989	8.8
Applied Sciences-Basel	92	3.008	2.8
Expert Systems with Applications	57	1.863	8.5
Neurocomputing	55	1.798	6.0
Sustainability	55	1.798	3.9
Applied Intelligence	53	1.733	5.3
Journal of Web Semantics	51	1.667	2.5
Semantic Web	50	1.635	3.1
IEEE Transactions on Knowledge and Data Engineering	38	1.242	8.9

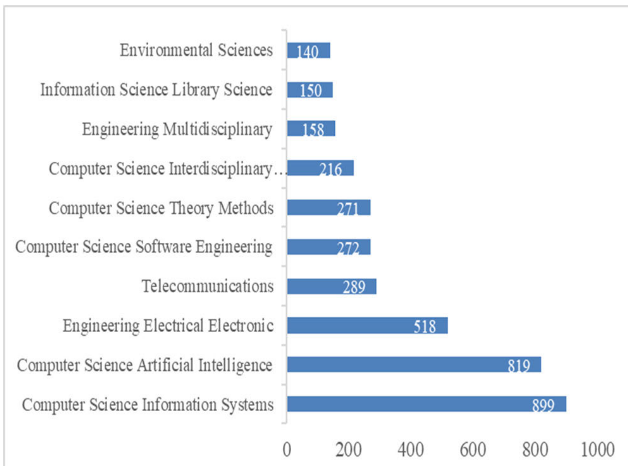

FIGURE 3. Top 10 subject categorie (2013-2022).

D. AUTHORS

The retrieved data indicated that up to 9,561 authors took part in the domain of knowledge graph research. The top 10 prolific authors who contributed over nine articles regarding knowledge graphs are shown in Table 5, of which six are from Chinese institutions. To be more specific, as for their affiliations, four authors are from National University of Defense Technology; two authors come from

Southeast University. At the same time, in terms of published articles, Lehmann ranked the top in the ranking list with 17 articles, and he mainly concentrated on the analysis of semantic question-answering systems and link prediction of knowledge graphs [26], [27]. Fernández took second place with 12 publications, which mainly specialized in the application of knowledge graph embedding models and natural language processing in the domain of biomedical literature [28]. What's more, Tan and the other two authors tied for third place with 11 articles, and he investigated generic knowledge graph embedding models and knowledge graph representation because his most contributed topics were domain relation extraction and relational knowledge prediction [29], [30]. Furthermore, Kraft from the United Kingdom and Recupero from Italy were also prolific authors. For instance, Recupero has nine publications related to the research of novel knowledge graph generation approach and the internal mechanism of deep learning transformer patterns for entity extraction as well as relations [31].

TABLE 5. Top 10 prolific authors with over nine publications (2013-2022).

Authors	Institutions	Publications
Jens, Lehmann	University of Bremen, Germany	17
Daniel, Domingo-Fernńdez	University of Bonn, Germany	12
Zhen, Tan	National University of Defense Technology, China	11
Guilin, Qi	Southeast University, China	11
Markus, Kraft	University of Cambridge, United Kingdom	11
Weidong, Xiao	National University of Defense Technology, China	10
Meng, Wang	Southeast University, China	10
Xiang, Zhao	National University of Defense Technology, China	10
Weixing, Zeng	National University of Defense Technology, China	9
Diego Reforgiato, Recupero	University of Cagliari, Italy	9

What's more, we employed VOSviewer to generate the co-authorship network map, which is demonstrated in Fig. 4. As an author's minimum document number was set at seven, 68 authors in total were finally confirmed to reach the threshold. The generated network map is composed of 12 nodes and 25 links, which belonged to three clusters respectively. To be specific, each node represents an individual author in the map. The connection between two nodes refers to the relation among co-authors, and the thickness rests with the outputs of co-authored publications [32]. The colors of the three clusters imply different co-author groups and their relations. Take the green cluster as an example, it was made up of four authors, and their relatedness is much closer than that of other authors in the other two clusters. Likewise, the red cluster is constituted by six co-authors, among which Xiang Zhao is taken as the pivot in the network. His closest co-author was Zhen Tan as the connection thickness between them reaches eight (Link strength=8). Driven by similar research interests, most of Zhao's articles took Zhen Tan as his close collaborator in embedding models by means of entity rotations and dynamic relation spaces [29], [33]. What's more, it should be noted that four out of the productive authors during the examined period occurred in this map, which further proves co-authorship could be a contributing element that promoted publication output.

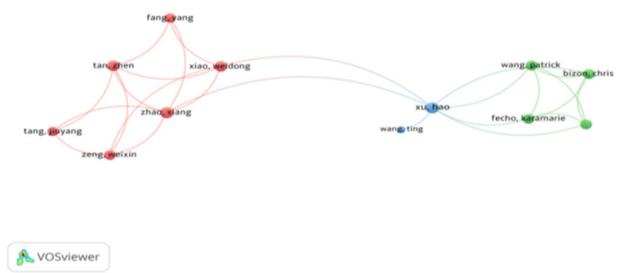

FIGURE 4. The co-authorship network map (2013-2022).

E. KEYWORDS

Being an essential index, keywords can directly unfold the key aspects of research articles, which play a significant part in decoding the foci as well as future directions of a specific discipline field [34]. Fig. 5 presents the keyword co-occurrence network map of knowledge graphs by the VOSviewer software. We took 18 as the minimum threshold for keyword occurrence, and the top 50 author keywords were eventually counted. It can be found that the node dimension in the map implies its frequency of occurrence. In other words, a larger node implies a higher frequency of occurrence [34]. According to the co-occurrence network of author keywords, the nodes and labels of knowledge graph, ontology and knowledge graph embedding are the most prominent. Specifically, the frequently occurred keywords were "knowledge graph(s)" (1,065 occurrences), "ontology and ontologies"(238), "deep learning"(129), "task analysis"(125), "semantics"(116), "knowledge graph embedding"(100), "semantic web"(96), "web"(93, "knowledge representation"(90) and "knowledge engineer"(85).

Furthermore, the relation strength between two nodes means the frequency of co-occurrence of a pair of nodes [35]. Take the node "knowledge graph" as an example, it has stronger relations with the nodes of "ontology", "semantic web" and "deep learning" within 57 links, which displays that these sets of keywords might be inclined to appear in the same article. In the meantime, as a given cluster is composed of nodes with the same color, the network map manifests nodes belonging to the same cluster are likely to occur in the studied articles of knowledge graphs. Clearly, these keywords can be classified into five thematic clusters, including ontology (red), relation extraction (yellow), knowledge graph embedding (blue), knowledge representation (green) and machine learning (purple). The evidence from node clusters indicates the author keywords are closely related with several themes, referring to ontology modelling, knowledge extraction, graph-based knowledge representation, knowledge graph embedding and applications of knowledge graphs.

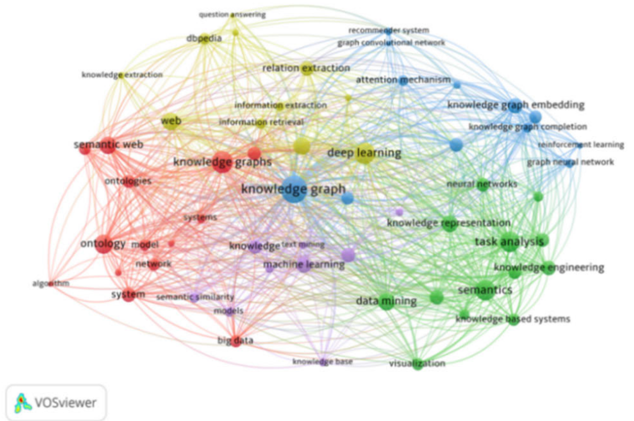

FIGURE 5. The co-occurrence network of top 50 author keyword.

To detect diachronic changes and trends of author keywords, Fig. 6 demonstrates the visualized distribution of author keywords in terms of the annual co-occurrence of keywords. During a given research span, the nodes of the keywords map were marked with different colors accordingly. For instance, the keywords with dark green indicate the earlier occurrence [36], while the keywords with light yellow mean the later occurrence. The observed publications in the past three years have focused on the following terms, such as ''graph neural network'' (avg. pub. year=2021.74, occurrences=50), ''knowledge graph completion'' (2021.12, 74), and ''knowledge based systems'' (2021.26, 41).

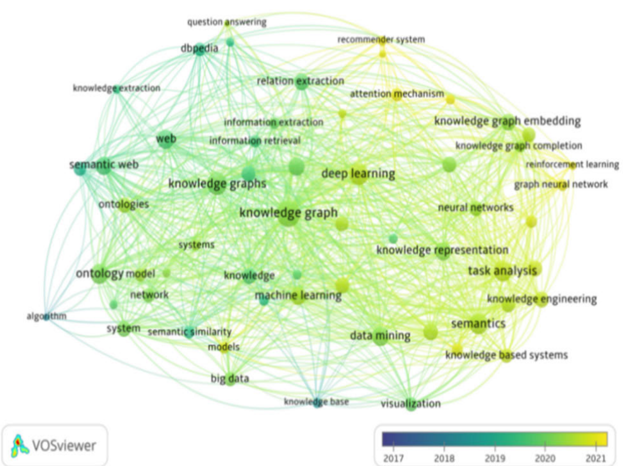

FIGURE 6. Annual distribution of top 50 author keywords.

IV. DISCUSSION

As for this study, the visualized data of the studied academic articles provided a macro overview of the newest development and discovered the emerging themes in knowledge graph research over the past decade. With the aid of the statistical analysis of author keywords co-occurrence, the frequently discussed hot topics in this field are summarized below.

A. ONTOLOGY MODELLING IN KNOWLEDGE GRAPHS

As stated by Gruber [37], ontology mainly refers to a formal and explicit account of common conceptualizations in the shape of concepts and relationships, and it can be thought of as the basic element of knowledge graphs. As it were, an ontology can be viewed as a fundamental framework for knowledge graphs. According to its architecture, knowledge graphs integrate information into ontologies and use inference engines to achieve the generation of new knowledge [38]. Combining ontologies with knowledge graphs is a reliable method of recognizing the complicated relationships between entities. General domain ontologies can act as a hub to allow the interconnection of specific information among different domains [39]. The previous research was focused on the approaches for ontology modelling in the knowledge graph environment, including logic-based models, structural-based models and hybrid models. In addition, many domain knowledge graphs have been constructed based on the ontology constraints and the mapping between the ontology models and the specific graph database.

B. KNOWLEDGE EXTRACTION FOR KNOWLEDGE GRAPH CONSTRUCTION

Since the term extraction and its related terms of information extraction and feature extraction are regularly used in the studied articles, knowledge extraction can be identified as a popular topic within the studied period. As the term implies, knowledge extraction is a kind of technique used for extracting information from unstructured or semi-structured data in an automatic or semi-automatic manner, which mainly takes relation extraction and named entity extraction as research foci [40]. Relation extraction is to extract unknown semantic relations between entities from multi-domain raw data and construct knowledge webs employing the mapping of semantic relations. A number of supervised or semi-supervised methods are utilized during the processes of relation extraction [7]. Compared with the former, named entity extraction is a process of performing entity recognition from data by rule-based methods, learning-based algorithms and neural network inferring systems. The completeness and accuracy of entity recognition have a direct impact on the quality of knowledge graphs. For example, Al-moslmi et al. [41] presented an overview of advances in extracting the named entities in the text, focusing on entity disambiguation as well as entity linking. At present, knowledge extraction is faced with problems that remain to be addressed, cross-language and open-domain knowledge extraction in particular.

C. EMBEDDING ISSUES IN KNOWLEDGE GRAPHS

As terms and phrases of embedding algorithms and neutral networks frequently occur in the articles of knowledge graphs, embedding issues have been studied across the examined period. Knowledge graph embedding, also known as knowledge representation learning, refers to a process of embedding components of a knowledge graph into continuous vector spaces [5], serving as a supporting role in measuring the correlation of entities and relations. Its capacity in semantic encoding in vectors and the performance of application service contribute to the gradual popularity and extensive use of knowledge graph embedding. Commonly, models of knowledge graph embedding fall into two categories: link prediction models for a single knowledge graph and entity alignment models for multiple knowledge graphs. Researchers addressed the processing flow of embedding models and compared the similarities and differences in the aspects of implementation methods and semantic-capturing levels. Some typical embedding models were constructed successively, such as TransE and MTransE. Innovative embedding techniques have been proposed for achieving success on relational learning tasks. Gao et al. [42] developed a novel knowledge embedding model by using the triple context of each triple, which proved critical to achieving sustained improvements. Given the significance of embedding issues in knowledge graphs, more insightful studies are expected to appear in this dimension.

D. GRAPH-BASED KNOWLEDGE REPRESENTATION

The term representation and its relevant term of knowledge engineering repeatedly appear in the studied articles, showing the growing interest of researchers in this topic. Knowledge representation could be considered the basis of knowledge graph construction and completion [43], which is mainly devoted to the study of how an enormous amount of knowledge can be represented in a computer-processable form. It uses a resource description framework (RDF) triples to describe the relations between entities, aiming to support the storage and update of knowledge. Recent advances in deep learning stimulate the emergence of developing knowledge representation approaches, which have aroused much attention from researchers in this field. Moreover, Wang et al. [44] developed a fresh learning method for text enhancement knowledge representation, and an interactive attention mechanism was employed to upgrade the accuracy of textual representations. Meanwhile, with the advancements in AI and the expressivity of knowledge graphs, knowledge representation learning oriented for entities and relations in knowledge bases has also gained popularity. Moreover, it is noted that graph-based knowledge representation incorporating the spatio-temporal dimension has been an important topic.

E. MULTI-MODAL KNOWLEDGE GRAPHS

With the development of digital information, a large number of multi-modal resources, such as images, videos, and audio, have emerged. Meanwhile, multi-modal knowledge graphs are becoming increasingly important as they can integrate multiple modalities into a single graph, providing a comprehensive representation of complex data [45]. Based on textual relationship triples, multi-modal knowledge graphs introduce multi-modal information into knowledge graphs and construct cross-modal entities and semantic relationships. Thus, knowledge graphs that merge heterogeneous signals enable complex reasoning and query operations. Notably, advances have been made in technical difficulties and critical tasks of multi-modal knowledge graphs, involving entity alignment, link prediction, fusion strategies, and the like. For instance, Zhu et al. [46] designed a novel framework named DFMKE, a dual fusion multi-modal knowledge graph embedding framework, for solving entity alignment between pairs of knowledge graphs, and experimentally demonstrated its performance of the new approach. Likewise, Wilcke et al. [47] devised a multi-modal message passing network that allows end-to-end learning not only from the graphs' structure, but also from the set of multi-modal node features, where the encoders are used to learn embeddings for nodes features of multiple modalities. Their study demonstrated the possibility and performance of combining knowledge graphs with multi-modal learning. Moreover, Zhang et al. [48] proposed a new model-agnostic multi-modal analogical reasoning framework with Transformer utilizing the structure mapping theory, which improves the traditional setting of analogical learning and opens up new approaches for enhancing analogical reasoning via multi-modal resources. Overall, there is much room for development and innovation in multi-modal knowledge graph research, especially in technical methods and the construction of graph data.

F. KNOWLEDGE-AWARE APPLICATIONS

Terms related to machine learning and textual mining were repeatedly used across the studied period, which indicates the trend of technology application in knowledge graph studies. Knowledge-aware applications mainly refer to achieving knowledge graph-oriented intelligent services in combination with specific scenarios through searches and recommendations. In terms of application areas, industry, agriculture, medicine and education have all benefited from the potential of knowledge graphs. For instance, knowledge graph technology brings the possibility of solving the problems of correlation representation and relevance searching and reasoning for data and knowledge in the manufacturing domain, so it plays an increasingly important role in the realization of intelligent manufacturing. In addition, these applications can help users select more targeted resources and provide scope for the realization of intelligent services, such as personalized recommendations, intelligent question answering (IQA), and the like. Related studies have shown that graph-based recommender systems and IQA systems have attracted great attention in this field. To enhance the quality of recommendation systems, multi-modal graph attention techniques were adopted to carry out information dissemination as well as recommend relevant resources [49]. Following users' personalized needs, different applications can support knowledge processing and generate knowledge interpretations from a variety of graphs. For example, based on domain-specific knowledge graphs, IQA systems with an intuitive visualization were developed via constructing graph data query statements and accurate knowledge searches, which provided feasible solutions for responding to more questions automatically [50]. Although rich structured knowledge can be useful for intelligent applications, it remains a challenge to integrate knowledge graphs into computational frameworks of real-world applications.

Based on the emerging topics discussed above, this analysis provides the following insights for future work: (1) More breakthroughs in knowledge graph core techniques are needed. (2) The effective platform of domain knowledge graphs needs to be constructed. (3) The progression from knowledge graphs to cognitive graphs might be advanced. Specifically speaking, first of all, some bottlenecks in knowledge graph core techniques are expected to break through, such as techniques for obtaining the relations for relation extraction and methods for resolving ambiguities in multiple source heterogeneity, to name a few. These difficulties are the challenges for expanding the scope of applications in the area of knowledge graph. Then, as the practical applications of knowledge graphs extend from general knowledge graphs to domain knowledge graphs, it is of great need to build up an effective platform that facilitates the construction of domain-related knowledge graphs. Lastly, due to the significance of cognitive technology in AI, there is a need for researchers to explore how to construct high-quality cognitive graphs and focus on application cases that combine intelligent reasoning, cognitive intelligence and knowledge graphs.

V. CONCLUSION

Overall, knowledge graph research is currently in a stage of rapid development. An increasing interest in the area of knowledge graph has grown over the past decade. Production output of knowledge graphs experienced strong growth during the past four years. Knowledge graph research is mainly centered on the fields of AI, engineering and technology. Asian countries have led the way with a significant number of publications and sufficient financial support. According to the analysis of authors, most of the productive authors came from China that made great contributions to this subject during the past ten years. According to the frequency of author keyword co-occurrence, the frequently used keywords were "knowledge graph", "ontology" and "knowledge graph embedding". The findings could be conducive for researchers deepening the original perception and detecting new directions for future research. Finally, it should be noted that the above discussion mainly focuses on hot topics and trends of knowledge graph research in articles published in the last decade. However, the use of application software, such as software packages for constructing knowledge graphs, also needs to be given sufficient attention. Therefore, such open source software in knowledge graph research should be the focus of follow-up studies.

Like other related studies, there are limitations to this study. For one thing, it could be affected by the search query strategies. If the titles and abstracts of retrieved articles do not adopt any of the search formulas, a small number of articles might be missed unintentionally in the study. For another, it could be influenced by the adopted databases. This study only consisted of publications indexed in the WoS database. The retrieved articles from other authoritative data sources such as Scopus, EI Compendex and Chinese National Knowledge Infrastructure (CNKI) might be included in follow-up studies.

REFERENCES

[1] T. Keller and S. O. Tergan, "Visualizing knowledge and information: An introduction," in *Knowledge and Information Visualization* (Lecture Notes in Computer Science), vol. 3426. Berlin, Germany: Springer, 2005, pp. 1–23, doi: 10.1007/11510154_1.

[2] P. Chen, Y. Lu, V. W. Zheng, X. Chen, and B. Yang, "KnowEdu: A system to construct knowledge graph for education," *IEEE Access*, vol. 6, pp. 31553–31563, 2018, doi: 10.1109/ACCESS.2018.2839607.

[3] M. Kejriwal, "Knowledge graphs: A practical review of the research landscape," *Information*, vol. 13, no. 4, p. 161, Mar. 2022, doi: 10.3390/info13040161.

[4] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, pp. 489–508, Dec. 2016, doi: 10.3233/sw-160218.

[5] M. Rizun, "Knowledge graph application in education: A literature review," *Acta. Univ. Lodz*, vol. 3, no. 342, pp. 7–19, Aug. 2019, doi: 10.18778/0208-6018.342.01.

[6] Z. Chen, Y. Wang, B. Zhao, J. Cheng, X. Zhao, and Z. Duan, "Knowledge graph completion: A review," *IEEE Access*, vol. 8, pp. 192435–192456, 2020, doi: 10.1109/ACCESS.2020.3030076.

[7] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 494–514, Feb. 2022, doi: 10.1109/TNNLS.2021.3070843.

[8] J. Cui and S. Yu, "Fostering deeper learning in a flipped classroom: Effects of knowledge graphs versus concept maps," *Brit. J. Educ. Technol.*, vol. 50, no. 5, pp. 2308–2328, Sep. 2019, doi: 10.1111/bjet.12841.

[9] Z. Wu and F. Jia, "Construction and application of a major-specific knowledge graph based on big data in education," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 7, pp. 64–79, Apr. 2022, doi: 10.3991/ijet.v17i07.30405.

[10] Q. Wang, Z. Mao, B. Wang, and L. Guo, "Knowledge graph embedding: A survey of approaches and applications," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 12, pp. 2724–2743, Dec. 2017, doi: 10.1109/TKDE.2017.2754499.

[11] J. Yan, C. Wang, W. Cheng, M. Gao, and A. Zhou, "A retrospective of knowledge graphs," *Frontiers Comput. Sci.*, vol. 12, no. 1, pp. 55–74, Feb. 2018, doi: 10.1007/s11704-016-5228-9.

[12] S. Issa, O. Adekunle, F. Hamdi, S. S. Cherfi, M. Dumontier, and A. Zaveri, "Knowledge graph completeness: A systematic literature review," *IEEE Access*, vol. 9, pp. 31322–31339, 2021, doi: 10.1109/ACCESS.2021.3056622.

[13] S. Tiwari, F. N. Al-Aswadi, and D. Gaurav, "Recent trends in knowledge graphs: Theory and practice," *Soft Comput.*, vol. 25, no. 13, pp. 8337–8355, Jul. 2021, doi: 10.1007/s00500-021-05756-8.

[14] B. Abu-Salih, "Domain-specific knowledge graphs: A survey," *J. Netw. Comput. Appl.*, vol. 185, Jul. 2021, Art. no. 103076, doi: 10.1016/j.jnca.2021.103076.

[15] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich, "A review of relational machine learning for knowledge graphs," *Proc. IEEE*, vol. 104, no. 1, pp. 11–33, Jan. 2016, doi: 10.1109/JPROC.2015.2483592.

[16] X. Li and L. Lei, "A bibliometric analysis of topic modelling studies (2000–2017)," *J. Inf. Sci.*, vol. 47, no. 2, pp. 161–175, Apr. 2021, doi: 10.1177/0165551519877049.

[17] S. Liu and S. Zhang, "A bibliometric analysis of computer-assisted English learning from 2001 to 2020," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 14, p. 53, Jul. 2021, doi: 10.3991/ijet.v16i14.24151.

[18] A. Kalantari, A. Kamsin, H. S. Kamaruddin, N. Ale Ebrahim, A. Gani, A. Ebrahimi, and S. Shamshirband, "A bibliometric approach to tracking big data research trends," *J. Big Data*, vol. 4, no. 1, p. 30, Sep. 2017, doi: 10.1186/s40537-017-0088-1.

[19] L. Yang, Z. Chen, T. Liu, Z. Gong, Y. Yu, and J. Wang, "Global trends of solid waste research from 1997 to 2011 by using bibliometric analysis," *Scientometrics*, vol. 96, no. 1, pp. 133–146, Jul. 2013, doi: 10.1007/s11192-012-0911-6.

[20] S. N. A. Majid and A. R. Salam, "A systematic review of augmented reality applications in language learning," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 10, p. 18, May 2021, doi: 10.3991/ijet.v16i10.17273.

[21] A.-M. Fernández-Luque, M.-S. Ramírez-Montoya, and J.-A. Cordón-García, "Training in digital competencies for health professionals: Systematic mapping (2015–2019)," *El Profesional Inf.*, vol. 30, no. 2, Mar. 2021, Art. no. e300213, doi: 10.3145/epi.2021.mar.13.

[22] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010, doi: 10.1007/s11192-009-0146-3.

[23] L. Yang, T. Sun, and Y. Liu, "A bibliometric investigation of flipped classroom research during 2000–2015," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 6, p. 178, Jun. 2017, doi: 10.3991/ijet.v12i06.7095.

[24] C. Huan and X. Guan, "Sketching landscapes in discourse analysis (1978–2018): A bibliometric study," *Discourse Stud.*, vol. 22, no. 6, pp. 697–719, Jun. 2020, doi: 10.1177/1461445620928814.

[25] S. Caliskan, A. V. Korzhuev, Y. B. Ikrennikova, S. V. Efimushkina, L. Z. Karavanova, and A. R. Masalimova, "Computer's place in teaching and learning for university students in the web of science database," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 19, p. 166, Oct. 2021, doi: 10.3991/ijet.v16i19.26057.

[26] H. Zafar, M. Dubey, J. Lehmann, and E. Demidova, "IQA: Interactive query construction in semantic question answering systems," *J. Web Semantics*, vol. 64, Oct. 2020, Art. no. 100586, doi: 10.1016/j.websem.2020.100586.

[27] C. Xu, M. Nayyeri, Y.-Y. Chen, and J. Lehmann, "Geometric algebra based embeddings for static and temporal knowledge graph completion," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 4838–4851, May 2023, doi: 10.1109/TKDE.2022.3151435.

[28] H. Balabin, C. T. Hoyt, C. Birkenbihl, B. M. Gyori, J. Bachman, A. T. Kodamullil, P. G. Plöger, M. Hofmann-Apitius, and D. Domingo-Fernández, "STonKGs: A sophisticated transformer trained on biomedical text and knowledge graphs," *Bioinformatics*, vol. 38, no. 6, pp. 1648–1656, Mar. 2022, doi: 10.1093/bioinformatics/btac001.

[29] Z. Tan, X. Zhao, Y. Fang, and W. Xiao, "GTrans: Generic knowledge graph embedding via multi-state entities and dynamic relation spaces," *IEEE Access*, vol. 6, pp. 8232–8244, 2018, doi: 10.1109/ACCESS.2018.2797876.

[30] Z. Tan, X. Zhao, Y. Fang, B. Ge, and W. Xiao, "Knowledge graph representation via similarity-based embedding," *Sci. Program.*, vol. 2018, Jul. 2018, Art. no. 6325635, doi: 10.1155/2018/6325635.

[31] D. Dessí, F. Osborne, D. R. Recupero, D. Buscaldi, and E. Motta, "SCICERO: A deep learning and NLP approach for generating scientific knowledge graphs in the computer science domain," *Knowl.-Based Syst.*, vol. 258, Dec. 2022, Art. no. 109945, doi: 10.1016/j.knosys.2022.109945.

[32] N. Li, J. Kramer, P. Gordon, and A. Agogino, "Co-author network analysis of human-centered design for development," *Design Sci.*, vol. 4, Apr. 2018, Art. no. e10, doi: 10.1017/dsj.2018.1.

[33] X. Huang, J. Tang, Z. Tan, W. Zeng, J. Wang, and X. Zhao, "Knowledge graph embedding by relational and entity rotation," *Knowl.-Based Syst.*, vol. 229, Oct. 2021, Art. no. 107310, doi: 10.1016/j.knosys.2021.107310.

[34] X. Chen, J. Chen, D. Wu, Y. Xie, and J. Li, "Mapping the research trends by co-word analysis based on keywords from funded project," *Proc. Comput. Sci.*, vol. 91, pp. 547–555, 2016, doi: 10.1016/j.procs.2016.07.140.

[35] H. Liao, M. Tang, L. Luo, C. Li, F. Chiclana, and X.-J. Zeng, "A bibliometric analysis and visualization of medical big data research," *Sustainability*, vol. 10, no. 2, p. 166, Jan. 2018, doi: 10.3390/su10010166.

[36] W. Xu, L. Feng, and J. Ma, "Understanding the domain of driving distraction with knowledge graphs," *PLoS One*, vol. 17, no. 12, Dec. 2022, Art. no. e0278822, doi: 10.1371/journal.pone.0278822.

[37] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowl. Acquisition*, vol. 5, no. 2, pp. 199–220, Jun. 1993, doi: 10.1006/knac.1993.1008.

[38] S. Auer, V. Kovtun, M. Prinz, A. Kasprzik, M. Stocker, and M. E. Vidal, "Towards a knowledge graph for science," in *Proc. 8th Int. Conf. Web Intell., Mining Semantics*, Jun. 2018, pp. 1–6, doi: 10.1145/3227609.3227689.

[39] S. Ferilli and D. Redavid, "An ontology and knowledge graph infrastructure for digital library knowledge representation," in *Digital Libraries: The Era of Big Data and Data Science* (Communications in Computer and Information Science), vol. 1177. Cham, Switzerland: Springer, Jan. 2020, pp. 47–61, doi: 10.1007/978-3-030-39905-4_6.

[40] Y. L. Zhang and B. Zhao, "Evolution of development trends and research hotspots of knowledge graph at home and abroad," *Library Theory Pract.*, vol. 4, pp. 121–128, Aug. 2021, doi: 10.14064/j.cnki.issn1005-8214.2021.04.019.

[41] T. Al-Moslmi, M. G. Ocaña, A. L. Opdahl, and C. Veres, "Named entity extraction for knowledge graphs: A literature overview," *IEEE Access*, vol. 8, pp. 32862–32881, 2020, doi: 10.1109/ACCESS.2020.2973928.

[42] H. Gao, J. Shi, G. Qi, and M. Wang, "Triple context-based knowledge graph embedding," *IEEE Access*, vol. 6, pp. 58978–58989, 2018, doi: 10.1109/ACCESS.2018.2875066.

[43] J. Lin, Y. Zhao, W. Huang, C. Liu, and H. Pu, "Domain knowledge graph-based research progress of knowledge representation," *Neural Comput. Appl.*, vol. 33, no. 2, pp. 681–690, Jan. 2021, doi: 10.1007/s00521-020-05057-5.

[44] Y. Wang, H. Zhang, G. Shi, Z. Liu, and Q. Zhou, "A model of text-enhanced knowledge graph representation learning with mutual attention," *IEEE Access*, vol. 8, pp. 52895–52905, 2020, doi: 10.1109/ACCESS.2020.2981212.

[45] Y. Chen, X. Ge, S. Yang, L. Hu, J. Li, and J. Zhang, "A survey on multimodal knowledge graphs: Construction, completion and applications," *Mathematics*, vol. 11, no. 8, p. 1815, Mar. 2023, doi: 10.3390/math11081815.

[46] J. Zhu, C. Huang, and P. De Meo, "DFMKE: A dual fusion multi-modal knowledge graph embedding framework for entity alignment," *Inf. Fusion*, vol. 90, pp. 111–119, Feb. 2023, doi: 10.1016/j.inffus.2022.09.012.

[47] W. X. Wilcke, P. Bloem, V. de Boer, and R. H. van t Veer, "End-to-end learning on multimodal knowledge graphs," 2023, *arXiv:2309.01169*.

[48] N. Zhang, L. Li, X. Chen, X. Liang, S. Deng, and H. Chen, "Multimodal analogical reasoning over knowledge graphs," 2022, *arXiv:2210.00312*.

[49] R. Sun, X. Cao, Y. Zhao, J. Wan, K. Zhou, F. Zhang, Z. Wang, and K. Zheng, "Multi-modal knowledge graphs for recommender systems," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2020, pp. 1405–1414, doi: 10.1145/3340531.3411947.

[50] Y. Tang, H. Han, X. Yu, J. Zhao, G. Liu, and L. Wei, "An intelligent question answering system based on power knowledge graph," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Washington, DC, USA, Jul. 2021, pp. 01–05, doi: 10.1109/PESGM46819.2021.9638018.

GANG WANG was born in Chaohu, China. He received the Ph.D. degree from the Lyceum of the Philippines University, in 2023. He is currently an Associate Professor with the School of Foreign Languages, Chaohu University, China. His main research interests include educational technology and second language acquisition.

JING HE was born in Hanshan, China. She received the M.A. degree from Anhui Normal University, China. She is currently an Associate Professor in applied linguistics with the School of Foreign Languages, Chaohu University, China. Her research interests include educational technology and corpus linguistics.