Mathematical Biosciences and Engineering

MBE, 20(2): 3731–3748.
DOI: 10.3934/mbe.2023175
Received: 19 October 2022
Revised: 22 November 2022
Accepted: 26 November 2022
Published: 09 December 2022

http://www.aimspress.com/journal/MBE

Research article

A vision sensing-enhanced knowledge graph inference method for a healthy operation index in higher education

Yu Nie*, Xingpeng Luo and Yanghang Yu

Yunnan University of Finance and Economics, Kunming 650221, Yunnan, China

* **Correspondence:** Email: zz2013@ynufe.edu.cn.

Abstract: We adopted the method of knowledge mapping to conduct in-depth visualization to propose the construction method of knowledge mapping-based inference of a healthy operation index in higher education (HOI-HE). For the first part, an improved named entity identification and relationship extraction method is developed, incorporating a vision sensing pre-training algorithm named BERT. For the second part, a multi-decision model-based knowledge graph is used to infer the HOI-HE score by using a multi-classifier ensemble learning approach. The combination of two parts constitutes a vision sensing-enhanced knowledge graph method. The functional modules of knowledge extraction, relational reasoning and triadic quality evaluation are integrated to provide the digital evaluation platform for the HOI-HE value. The vision sensing-enhanced knowledge inference method for the HOI-HE is able to exceed the benefit of pure data-driven methods. The experimental results in some simulated scenes show that the proposed knowledge inference method can work well in the evaluation of a HOI-HE, as well as to discover some latent risk.

Keywords: knowledge graph; vision sensing; healthy operation index; data visualization; education management

1. Introduction

Education is a great cause of cultivating people with high-value involvement, and educational management serving education, naturally, also has multiple missions, so the importance of educational management research is self-evident [1]. Visualization technology combines the theories and methods of statistics, applied mathematics, computer science, information science and bibliometrics, which can more conveniently and intuitively show the development history, research subjects, research hotspots, research bases and research frontiers of a certain research field [2]. By selecting research data on the topic of educational management and trying to use visual analysis software and methods to map the knowledge of educational management research in a certain period, we can help researchers "know the sequence" and grasp the research hotspots, the historical lineage and future trends of educational management [3]. It can help researchers to grasp academic groups and research hotspots, analyze the changing status of educational management research and explore research trends. By establishing its education quality monitoring data platform [4], it can use information technology to monitor some data that can objectively reflect the objective teaching status of colleges and universities, such as the status of teachers, the number of students, scientific research results, financial status, the number of teaching equipment and facilities, the employment situation of students after graduation, the number of school books, etc., to help colleges and universities accurately grasp the overall development of education [5]. It can also help universities to improve the shortcomings in education and teaching and effectively promote the internal development of higher education [6]. At the same time, it can help colleges and universities save labor and material costs, improve the efficiency of filling and reviewing and facilitate data management [7].

As an excellent technology for storage and application with highly structured knowledge organization, knowledge graphs extract and store the knowledge of entities, attributes and relationships existing in the massive heterogeneous data in the network by using natural language processing and other techniques [8]. Originating from the semantic web, a knowledge graph is proposed for the problem of incomplete information in traditional data retrieval, providing a query environment, which is a network composed of data. The knowledge graph is a product of artificial intelligence, which can provide users with an accurate and clear graph structure and help them quickly access domain information. With its powerful visualization capability, users will no longer be directly confronted with messy and disorganized information, unlike the previous presentation of traditional websites, but will be replaced by clear diagrams. The information attributes and connections of colleges and universities are more suitable to for storage in knowledge graphs. Collecting basic information about colleges and universities and obtaining valuable information via knowledge graph visualization can be studied, but there are relatively few studies applying knowledge graphs to the information display of colleges and universities. A more complete and standardized domain knowledge construction system has not been formed, and there is no mature knowledge graph construction method for college information [9]. Therefore, it is of certain practical significance to research the construction of the knowledge mapping of university information in this paper.

Knowledge graphs are presented to users in the form of diagrams that link different entities and their unique attributes together. Compared with traditional text-based websites, this relational network structure no longer directly confronts users with disordered information knowledge, and the graphs can better reveal the important content of university information. The data in the network grow greatly and are full of invalid links and a lot of advertisements, so it is not easy for users to find out useful information precisely. The construction of a university information knowledge map can not only effectively sort out the characteristic information of universities, but it can also show its rich hierarchical connections. An algorithm for using recommended words based on collaborative combination filtering is proposed. It greatly helps users to better understand the information of each university and gives some relevant reference opinions to decision-makers. The construction of a modern network resource platform helps users to quickly navigate the knowledge links between entities and provides important academic values to promote scientific and technological progress and knowledge system development in the education industry. Therefore, we selected representative profiles of universities as the main body and studied the construction scheme of the university information knowledge map. Then, the visualization model of the university information knowledge map was built and the visualization platform was designed for experimental display. The structural relationship, as well as the distribution of knowledge points, are presented intuitively, which makes a positive contribution to the research of knowledge mapping visualization. This scheme can be applied in subsequent research to gradually complete the university knowledge information and build a more comprehensive university knowledge map.

2. Related work

In terms of knowledge mapping applications, Ji et al. argued that the purpose of knowledge mapping is discovery, understanding, communication and education, which can be applied to disciplinary visualization [10]. Yang et al. argued that knowledge mapping can provide a comprehensive description of the rapidly developing field of arts and sciences, based on which a web-based knowledge mapping system was developed [11]. In a previous study, Wang and Lu summarized that knowledge mapping application development can be divided into two periods, i.e., the exploration period and the development period [12]. Decuypere and Landri used space-filling visual representation for gene data ontology visualization [13]. Node linkage graphs use circles to represent nodes, lines to represent edges and different colors to represent different categories, and force orientation is one of the common layouts for laying out node links. Other visualization methods are indented lists and Eulerian lists, but both methods are not commonly used and not intuitive enough [14]. Cerezo et al. composed and analyzed the literature published in educational management in the past 10 years, and they believe that the current research themes in educational management are mainly historical research, theoretical introduction and discipline construction [15]. This high-dimensional method of robot semantic learning is directly displayed by using the visual map technology associated with knowledge points, and all the dynamic semantic learning data that are common in the industry are directly embedded into a relatively low-dimensional dynamic semantic data space. The main problems are that there are more introductory things and fewer original things, not enough breakthroughs in methods and few collaborative studies.

In the approach of ontology construction and general knowledge mapping, Chang et al. extracted the entities, attributes and relationships among entities from data sources of different grades and different chapters to construct a knowledge map of mathematics subjects, through which they aimed to improve the problems of poor learning efficiency and difficulty with the concentration of students [16]. Luan and Tsai constructed the knowledge map of Tang poetry in the big data environment and provided intelligent services such as knowledge exploration and semantic query. [17]. The evaluation based on the gold standard has a mature application in ontology construction, which mainly compares the data with the existing excellent "gold standard", lists the shortcomings of its knowledge map and makes improvements [18]. The advantage of this method is that, based on the "gold standard", it can directly calculate the accuracy and recall rate and minimize unnecessary time.

In the study of knowledge mapping in basic education, it has been found that the construction or application of knowledge mapping for basic education disciplines is generally conducted only for certain disciplines, and the research objects are mostly mathematics, biology, chemistry and other disciplines with strong scientific characteristics, and there is a lack of research on disciplines with liberal arts; so, the scope of application of such mapping is too narrow. Therefore, we aimed to build a general knowledge mapping method in basic education to improve the application of knowledge mapping in basic education.

3. Visual analysis of knowledge graph for higher education management research

3.1. Knowledge map construction for education management

The collection and processing of data is the first step to doing any work, and it is only when the right step is taken will everything afterward become smoother. In this paper, we make general anticipation of knowledge mapping, combine the process and methods of creating knowledge maps of existing disciplines and decide to divide all of the data into the following parts. Various websites on the Internet contain a rich variety of learning resources, but searching for learning resources only through search engines may result in wasting a lot of time to find some content that does not match, and it is difficult to effectively help students learn relevant knowledge and improve relevant skills [19]. Therefore, obtaining rich resources from the Internet and extracting relevant knowledge is significant to help students improve their abilities and learn relevant knowledge efficiently and quickly.

This chapter is a summary of a rich variety of learning resources from various websites on the Internet, as obtained by combining crawler technology and web-page noise reduction technology. For the acquired Internet learning data, our model mainly utilizes the Baidu encyclopedia engineering education entries as the starting point, continuously extracts relevant entries for search and performs web data crawling work. The core idea of the representation-based model is that the representations of items and users are obtained by aggregating multiple sources of information, such as knowledge graphs, user-item interaction matrices and the content or attributes of the items. Finally, the generated user representation u and item representation v are made as an inner product to represent the probability of that user clicking on the item, as shown in Eq (1). The most used methods are as follows: reducing the number of pixels used when displaying primitives, clipping the number of node relationships, increasing physical pixels and reducing the complexity of primitives.

$$\hat{y}_{uv} = uv^T \tag{1}$$

Knowledge graphs extract knowledge from raw data or third-party databases using automatic or semi-automatic techniques. The level of human cognition is constantly changing, so the entire knowledge graph is constantly updated when it is constructed. Generally, there are two types of knowledge graphs: top-down and bottom-up. The bottom-up type extracts entities first and then builds an ontology schema, and it is mostly used for general knowledge graphs. The top-down type builds an ontology layer first and relies on ontology to add entities to the data layer, and it is mostly used for domain knowledge graphs. In the actual construction, depending on the specific situation, there is no limited standard for the construction process of a knowledge map, and the suitable method is selected according to the actual task-specific analysis. In the construction process of the university information knowledge graph, to avoid adding invalid data and ensuring the quality of the knowledge graph, we adopted the top-down construction method.

The input vector and output vector of the module are arithmetically added as the final output vector of the model. Because it is more convenient to optimize by modifying the input vector than reconstructing the entire output vector, this operation can make the network training easier [20]. Therefore, it leads to some shortcomings of the algorithm because the negative triad vocabulary composed via this method may be essentially positive such that the potential probability of correctness of the negative triad would be particularly high, making the score function of that negative triad take similar values to the score function of the positive triad and become unable to be distinguished. We use these part-of-speech combination rules to perform rule matching on adjacent word strings of compound words. As a result, it can lead to errors in the final computation results obtained, as shown in Figure 1.

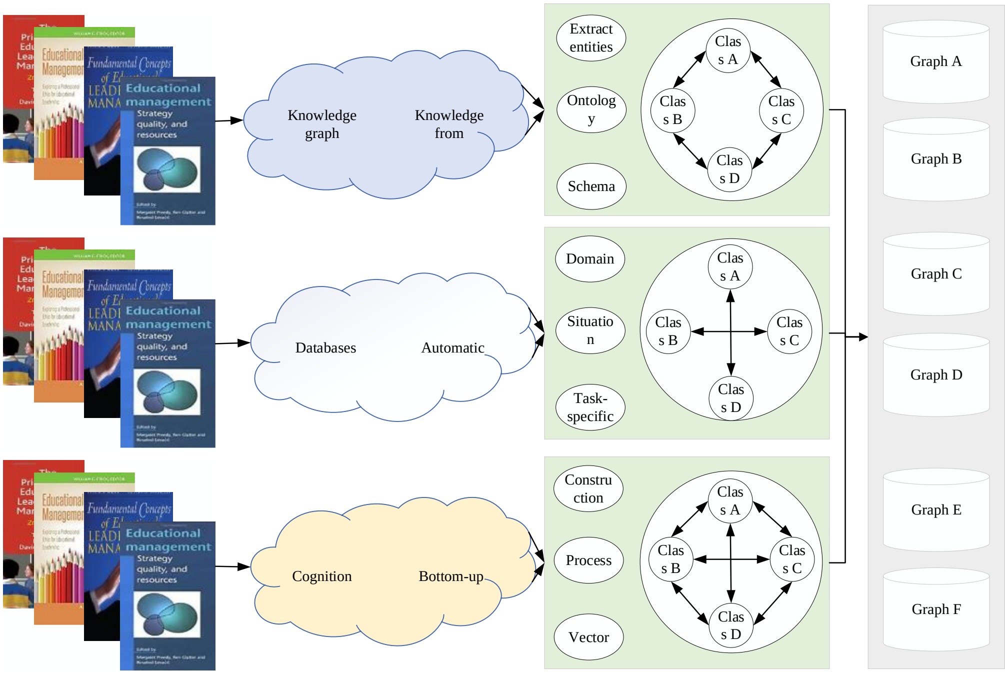

Figure 1. Conceptual diagram of visualization process.

The knowledge graph data contains a large amount of information about the realization paths of entity relationships that need to be supplemented, and the individual paths constituted by multiple relationships between two entities can describe the semantic relationships between these two entity vocabularies to a certain extent. If the rules are met, remove them; otherwise, leave them as candidate field compound words. The TransE improved probabilistic algorithm used in this work complements and improves the entity relationships that are missing in the originally constructed knowledge graph data by incorporating the information of relationship realization paths between triadic vocabularies into the vector representation of entities and relationships for calculation.

$$loss = \sum_{(u,i,j)\in O}^{\Sigma}\left(\ln\sigma\left(\hat{y}_{uv}(u,v)\right) + \hat{y}_{uv}(u,j) - \lambda\|\theta\|_2^2\right) \quad (2)$$

The TransE model continues to be trained by using an improved negative triad potential correct probability algorithm to obtain more accurate entity relationship results, which are used to complement and refine the triads in the knowledge graph. Today's knowledge graph visualization models are either full-image displays or provide text search boxes for a single interaction [21]. The full-image display cannot avoid high node density and serious knowledge occlusion, which causes visual confusion and weak reflection of the user interest level. The interaction method of search keywords is also flawed, and if users cannot clarify their search goals and query in the huge amount of data, they will have the problem of knowledge disorientation.

$$Confidence(L_i, L_j) = \frac{1}{L_i - L_j} - \frac{1}{L_i - R} - \frac{1}{L_j + R} \tag{3}$$

The knowledge graph varies from field to field, but its construction process is the same, mainly including the information extraction process and knowledge fusion process. An algorithm based on collaborative combinatorial filtering and the complexity that this algorithm only fully uses a matrix of item-service-customer combination evaluations without fully and carefully considering its usage semantics, etc., is proposed for the use of recommenders based on collaborative combinatorial filtering. This visual mapping technique of knowledge point associations is used to directly present this high-dimensional approach to robotic semantic learning, embedding all of the dynamic semantic learning data that are already prevalent in the industry into a lower-dimensional dynamic semantic data space [22]. By analyzing and computing all of the semantic combinatorial similarities between similar items, all of the semantic similarity information of each similar item itself is combined and integrated into the collaborative semantic filtering and combination recommendation. The algorithm features make up for the serious lack of the item collaborative combination filtering recommendation algorithm that does not fully consider the cultural connotation and related knowledge of the recommended items themselves, as well as greatly enhance the effectiveness of the item collaborative combination filtering algorithm recommendations at the level of comprehensive semantic analysis, as shown in Figure 2.

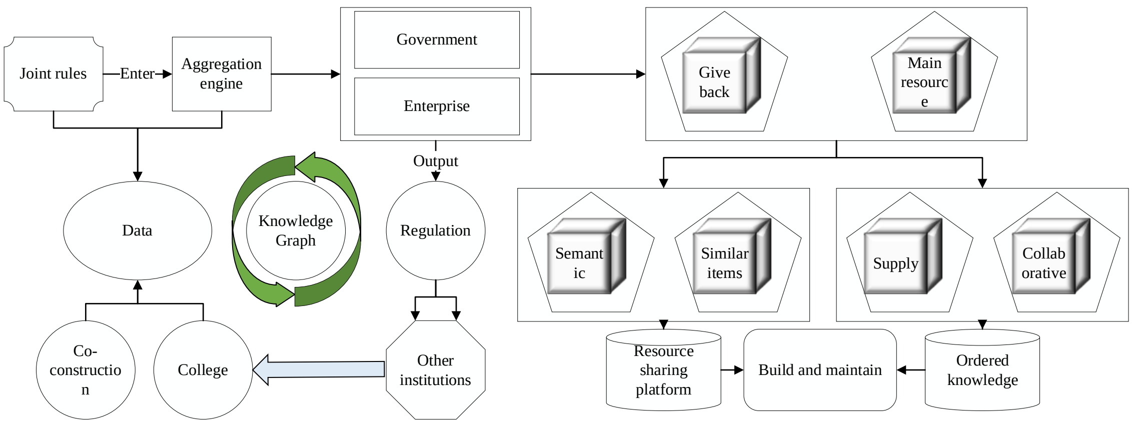

Figure 2. Production model of educational digital resources based on knowledge graph.

Relationships depend on entities to exist, and when there are too many entities, many relationships connect different entities to form a large-scale network graph. Users cannot intuitively discover the implied values in the relationship network, and to express as much value as possible using the limited space, the more commonly used methods are as follows: reducing the pixels used in the graph element display, cropping the number of node relationships, increasing the physical pixels and reducing the complexity of the graph element. Therefore, in this work, we designed a compact force-guided layout scheme when designing the relationship layout to reduce the complexity of the tuples by focusing on the user's interest, reducing—or removing—useless edges and making the nodes evenly distributed on the page.

$$ECII_i = W_i^2 I_i^2 \tag{4}$$

Nevertheless, some compound words with low word frequency will still be missed, because, in the process of multi-word combination, the probability of single-word words among them will be high, while the probability of compound words appearing may be low, so the value of their mutual information will be low. Then, there is the cooperation with the same content, which refers to the joint construction of the curriculum system and teaching content by schools and enterprises, which ensures the progressive development of engineering capabilities. In order to solve this problem, we added an anti rule vocabulary template based on mutual information for matching. These lexical combination rules are used to rule match the adjacent word strings of the compound words, and if they match the rules, they are removed; otherwise, they are left as candidate domain compound words.

According to the calculation of the index, which is an evaluation index for students to improve engineering ability, it is concluded that improving students' engineering ability not only needs the interplay of national training, school training, teacher training and student training, but also requires focus on starting from students' training so that students' engineering ability can be improved more effectively.

Moreover, student training should be combined with school training; for example, school-enterprise cooperation is an important way to train students to achieve engineering ability improvement. Through school-enterprise cooperation, the engineering ability of students can be improved in terms of practical operability so that the engineering ability of students can meet the needs of enterprises and lay a solid foundation for employment. Therefore, this work can carry out the task of school-enterprise cooperation from three aspects. The first is cooperation toward the same goal, with the goal of a solution being that everyone benefits for both schools and enterprises, consequently jointly formulating the training program of engineering innovation ability to ensure the consistency of the goal of school-enterprise training talents [23]. Corresponding to the entities obtained by the user interest model, the methods of navigation and selection were adopted, and the obtained results are displayed by means of mouse movement display and a click display method. Then, it is the cooperation with consistent content, which refers to the joint construction of curriculum system and teaching content between school and enterprise and ensures the progressive development of engineering ability. Finally, the cooperation of consistent evaluation refers to the joint development of evaluation indexes by schools and enterprises to stimulate students' innovation ability and enhance their practical ability.

3.2. Knowledge graph visualization design

After constructing the completed knowledge map of university information, the knowledge map data are used as input to solve the problems of visual confusion and knowledge lost in knowledge map visualization by reducing two dimensions of the complex structure of graph elements and interactive visual design [24].

Designing an interactive visualization scheme can remove the invalid edges. In consideration of the entities obtained from the user interest model, navigation and selection are performed, and the obtained results are displayed by using the mouse-over display and click display methods to show the relationships. Regarding the search model, the individual entities obtained by the search were designed in two dimensions: a refined display and a global display. All of the visualization view layers are displayed as node-link graphs, and the layout between two nodes, citing the D3.js library, is implemented by using the KK algorithm to achieve a uniform distribution. Users view university information under different query methods. Fuzzy query is applied to select conditions under the hierarchical structure.

When the number of nodes is large, it will result in large node density and the superposition of text information, making it difficult to mine useful values.

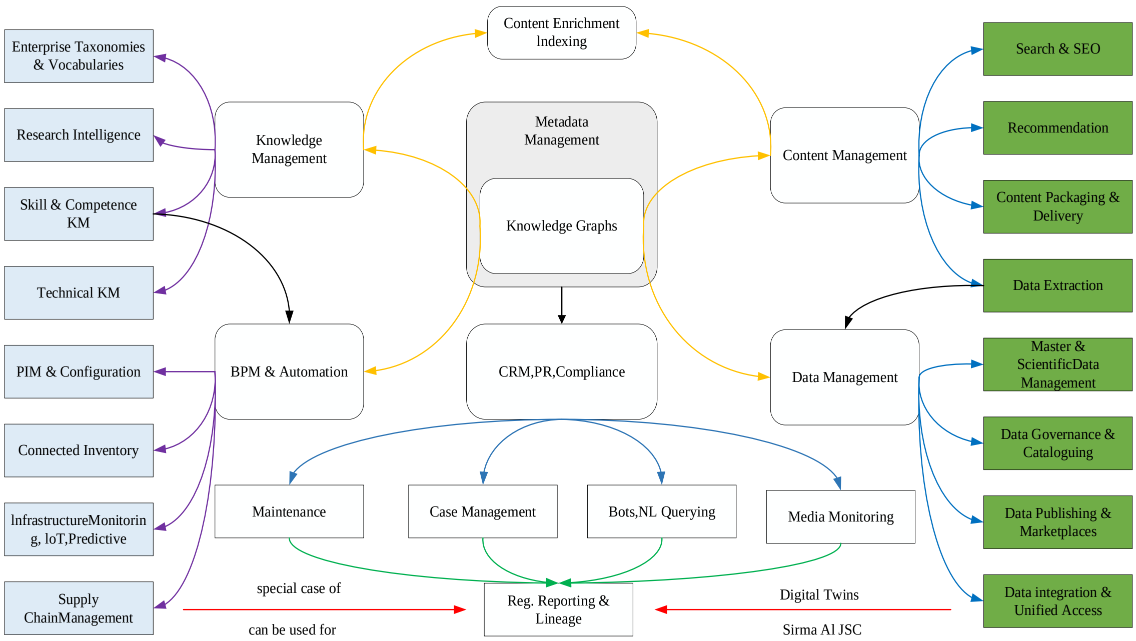

Figure 3. Knowledge graph visualization model.

As shown in Figure 3, for the nodes, the solution was adopted from the perspective of reducing the complex structure of graph elements, designing a user interest model, using ontologies that can be used as filtering conditions and allowing users to select the range of entities to view by themselves in combination with the degree of interest, which can effectively remove a part of information irrelevant to the target and alleviate problems such as knowledge loss. The user obtains a collection of university entities after filtering and then checks the university information one by one according to their own interests. In addition, an entity search model has been designed to precisely search the knowledge in which users are interested. For relationships, relationship visualization corresponds to entities, and there will be relationships when there are entities.

In this work, we used theoretical research and empirical analysis and read literature combined with relevant professional websites to sort out the construction methods of knowledge graph visualization models for university information. The present-day knowledge graph visualization models are either full-image displays or provide a text search box for a single interaction. The full-image display cannot avoid the high density of nodes and serious knowledge occlusion, which causes visual confusion and weak embodiment of user interest level. The interaction method of searching for keywords also has defects, and if users cannot clarify their search goals and query in the massive data, they will have the problem of knowledge loss. Therefore, we combined two perspectives of user interest and interaction to propose a knowledge graph visualization model based on university information.

Keyword-based data search and query methods are, to a certain extent, in line with the interests of actual users, and users can get results through such methods when they have clear search objectives. Compared with other search methods, the keyword-based search method can show the correlation between the search contents and the implicit semantic connection, which can reflect the user's expectation information at the semantic level. In addition, different users have different interest points, so the universities they want to view are also different. The way of hierarchical filtering alone is not enough for users to quickly focus their interests in a group of universities, so it is important to design a keyword search model for the development of interactive user visualization.

3.3. Design and implementation of interactive knowledge graph visualization

It is required to show the association structure between entities and entities in the knowledge map of university information, and to show it in the way of the node-link diagram, which can show both all and partial. Users can view college information under different query methods [25]. The fuzzy query is the selection of conditions under the hierarchical structure; users can get a collection of college entities after filtering with their interest level and then view college information one by one; a precise query is the search of specific college names directly through the search bar and then the viewing of college information. It provides various views, and the concise display highlights a few important nodes to quickly locate users' interest points so that users can choose whether to continue to click on college nodes to view their full knowledge map according to their interest level; the full display presents all of the contents associated with college nodes. The visualization view can support users to zoom, drag and drop, expand and collapse, in addition to allowing users to expand the hidden node content by clicking on the node.

Regarding the page functional area division module, this module is mainly purposed to assign the position of each area on the page to guarantee the effective operation of the knowledge graph visualization. There is a title bar area, node-link graph display area, navigation and selection area, search text box area, mode selection button area and text prompt comment area.

The user interest model module, which strictly adheres to the user interest, was designed to display the model. The main purpose of this part is to implement the filtering of entity nodes through the hierarchy. Based on the three available ontology concepts in the ontology, which are used as selection conditions, the user can get a part of the college entity nodes based on the corresponding constraints according to the interests. Based on the obtained results, a function of partial entity display with the paginated display is designed based on the node generation algorithm.

The keyword search model module is based on the already constructed entity search visualization model, which is a module mainly designed to realize the target entity through a precise search. Users enter the full name of the university in the text box according to their expected target and then search for the relevant results. According to the obtained results, a streamlined display structure that can only display the key information of universities is designed by researching and investigating users' interest level in various types of information of each university, and then users decide whether to view all information of universities according to their judgment, which is presented through all displays (as shown in Figure 4).

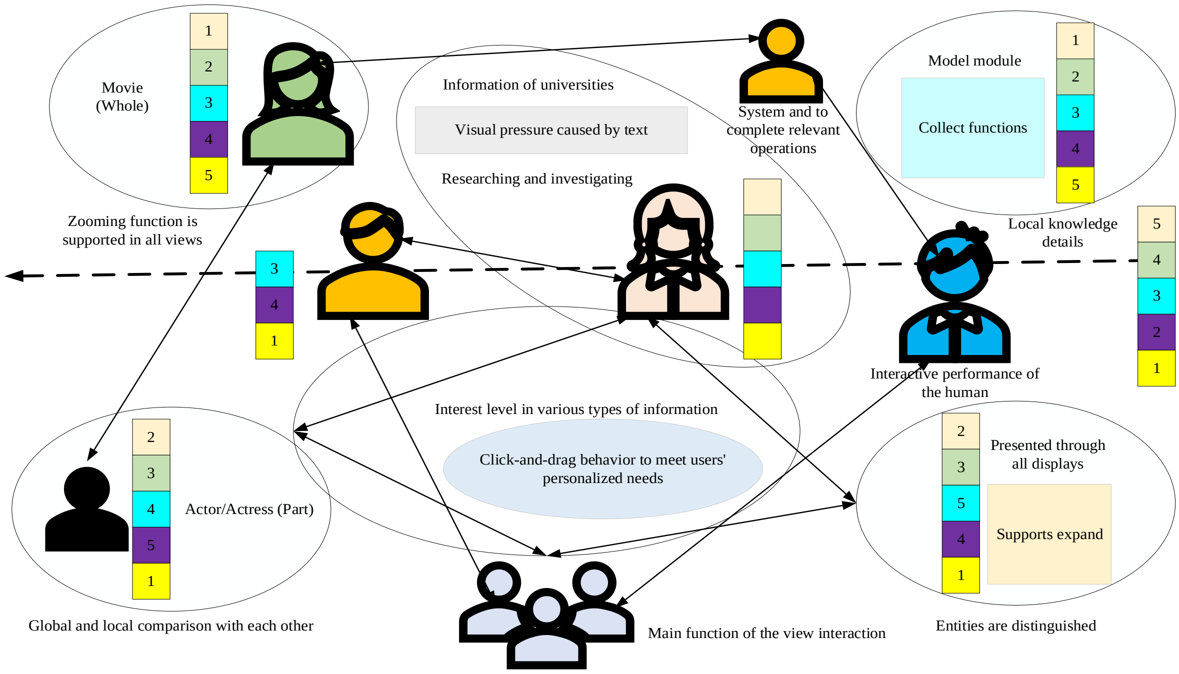

Figure 4. Interactive knowledge graph visualization system.

The main function of the view interaction design module is to reflect the interactive performance of the human and the system, and to complete relevant operations by using the mouse. The zooming function is supported in all views to view local knowledge details and to make a global and local comparison of each other. Different classes of entities are distinguished by node color and node size. The model also supports user click-and-drag behavior to meet users' personalized needs. Supports expand and collect functions, which can effectively relieve the visual pressure caused by text in the diagram and reduce confusion. When the mouse hovers over an entity, other information connected to the entity is displayed to provide users with the ability to view other entities around them.

4. Results and analysis

4.1. Results of knowledge graph construction for educational management

Educational organizations usually need a team of teachers with excellent skills and high-quality research capabilities in their respective professional fields. Each trainee may be a professional elite with strong learning skills from all over the country.

As a requester of educational resources, one needs to search for their keywords through the keyword and digital resource retrieval system in the blockchain network, such as keyword and word retrieval, find the desired resource, obtain the public key address of the owner of the keyword and digital resource and the unique representation information of one of the digital resources, and then send a request to the blockchain network through these two pieces of information, as well as the blockchain network. After verifying the identity of the keyword and the digital resource requestor and confirming whether the digital resource requestor already has the real request authority of the digital resource, the blockchain network will directly return a real resource address to the resource requestor, which is usually a URL or an address that can directly point to a real file. However, this address is not directly returned to the resource requester, as it is directly returned to the data obtained after the public key of the resource requester is encrypted in the email, and the real resource requester can directly decrypt their data and get the real email address if they use the private key in the email and then use this address to get the resource. The transaction of resources can be free or paid.

Among them, the process of requesting resources is also regarded as a transaction by the blockchain network, and this transaction record will also be recorded in the block permanently in the future on the blockchain network. The blockchain network can easily setup a virtual coin trading system, and the compilation and operation of smart contracts also require virtual coins. Virtual coins can be obtained through resource uploading and mining, as shown in Figure 5. The blockchain network can easily establish a virtual currency trading system, and the compilation and operation of smart contracts also require virtual currency support.

Due to many nodes and connections, the cluster names were changed to term clusters of the cited literature to better reflect the knowledge base revealed by the highly cited literature, and another knowledge map of the co-cited literature shown by the cluster labels was drawn. To accurately reflect the characteristics of highly cited literature in education management in China, we classified the data statistically and extracted the top 50 classical literature in the network.

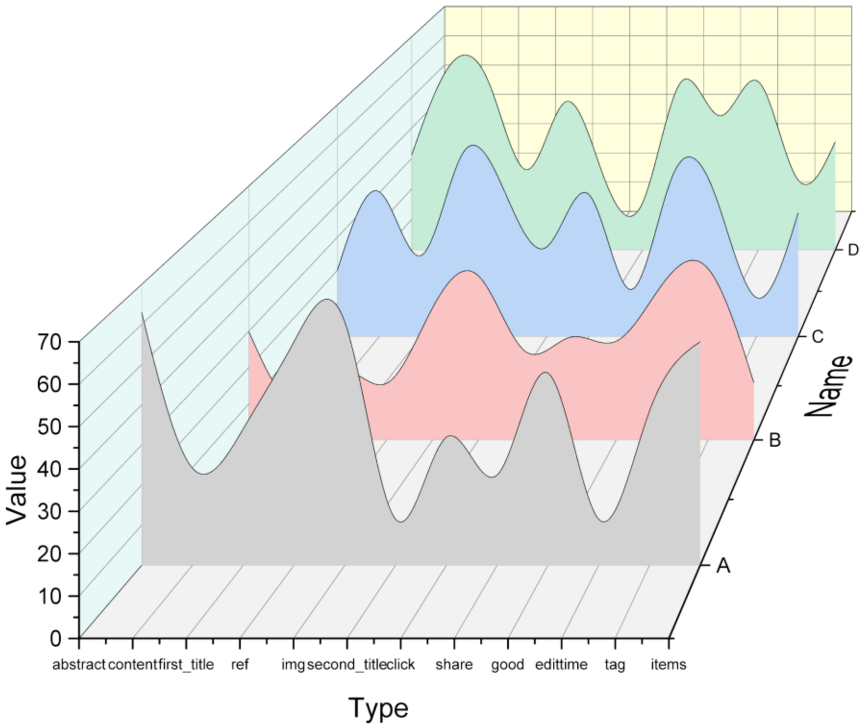

Figure 5. Some of the words' attribute data.

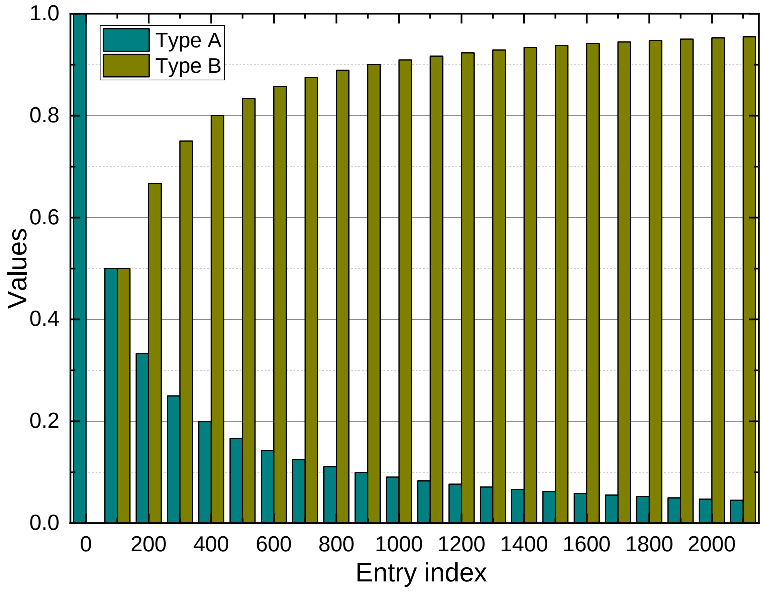

Figure 6. Multi-decision model's distribution of ratings for quality words.

In classification mode, the number of votes for each classification in the multi-decision vector is counted; also, the votes are weighted with the weights of the corresponding attributes to obtain the quality classification with the highest number of votes as the predicted value before being compared with the actual value to obtain the evaluation value of each classification. The dot product of the decision vector and the weight vector are used as the word score in scoring mode, and statistical distribution verification and regression verification are performed. In this study, 3200 words to be tested were scored by the multi-decision model, including 810 featured words and 2390 common words; the scores were sorted from highest to lowest, and the statistics were classified according to the labels of the test data. As shown in Figure 6, the scores of common words by the multi-decision model are generally lower than those of featured words, and the scores of words labeled as featured words are mostly above 2, while the scores of words labeled as common words are mostly 2.5 and below. This shows that the multi-decision model's scoring of words has a certain degree of differentiation in the quality of words.

By setting the regression segmentation threshold percentage, the words that have derived ratings can be regressed into the classification. In this study, the words with scores above the threshold are regressed to the high-quality category, the words with scores below the threshold are regressed to the low-quality category and the evaluation index after the regression of scores is calculated. "Score" denotes the score calculated from the regression segmentation threshold. In Figure 6, we can see that, as the threshold value increases, the regression accuracy decreases, the recall rate grows, the F1 value grows and then decreases and the F1-score reaches the global maximum when the percentage percent reaches 28%. Among them, there were 810 featured entries and 2390 common entries; the scores were sorted from high to low and classified and counted according to the labels of the test data.

The dynamic presentation of digital education resources is reflected in two aspects. First, according to the needs and knowledge levels of different learners, the content of learning resources will be dynamically adjusted and combined to better provide knowledge services to learners. The resources are kept in the latest and best condition.

If the classification mode is selected during the model validation, then, after the multi-decision vector is generated, it enters the voting and the traditional voting algorithm counts the most classification results in the multi-decision vector as the final classification result, e.g., there was a decision vector R = [1, 1, 1, 0, 0]; then, the final classification result was 1. And, the weighted voting method needed to weigh the votes to sum up, e.g., the weights were W = [0.1, 0.1, 0.05, 0.8, 0.2], the weighted final vote, Vote = [2, 0.75], and the final classification result was 0.

4.2. Visualization analysis results

The key core of educational digital resource service lies in the organization of resources, which refers to the classification and coding of multimodal educational digital resources and storing them to build an educational digital resource repository for easy organization, management and retrieval. Inspired by the library resource management mechanism, the organization of educational resources was designed to consist of three components: educational resource pointers, metadata and resource entities.

Among them, the pointer mainly preserves the continuity between the educational information resource entity and the description of the information resource. The metadatabase is mainly used to describe some form of description of educational information resources, while the educational digital resource entity is the resource itself. Knowledge management maps are widely used in the operation and management of human resources and labor and social security organizations in China. One of the early beginnings of knowledge maps is from what has been proposed by previous people that knowledge maps are detailed and interrelated nonlinear representations of thinking. The main purpose of the knowledge resource map is to guide and present the knowledge points of each discipline, and establish mutual relations. The knowledge resource map can clearly and accurately show the relationship between each discipline and knowledge resources. The so-called school knowledge data map mainly contains the following three important influencing factors. The first is the organization and construction of the relevant knowledge data ontology, the second is the organization and visualization of the relevant knowledge data ontology and the third is the organization and visualization of the relevant knowledge data of the school.

Literature co-citations are used to reflect the similarity of research contents of two documents by the number of co-citations in other literature, thus revealing some connections between the documents. To see the position of high-frequency literature more clearly in the network, as well as the connection between them, the bibliographic co-occurrence system was first used to generate a high-frequency literature co-occurrence matrix with a threshold value ≥ 9, and then the software was used to generate a knowledge map of high-frequency cited literature co-occurrence, as shown in Figure 7. This included the school's teaching staff, the number of students, scientific research results, financial status, the number of teaching equipment and facilities, the employment situation of students after graduation, the number of school books and other data to help colleges and universities accurately grasp the overall educational development status of the school to continuously improve education and teaching.

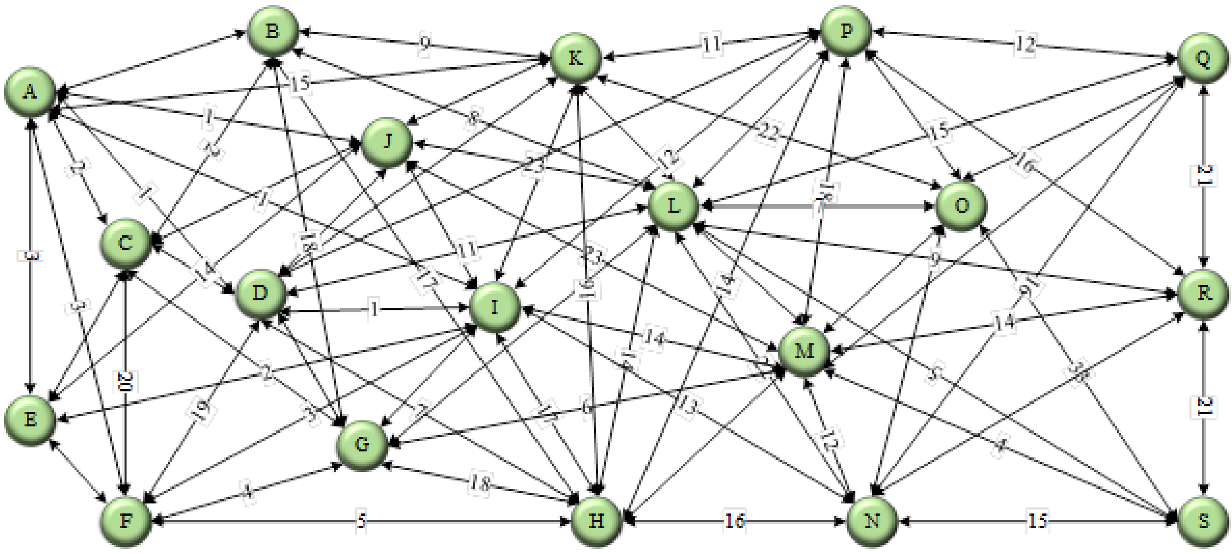

Figure 7. Visualization results of co-occurrence knowledge map of high-frequency cited literature.

The extraction of entity relations is the most important content of the whole knowledge graph, as the connections of knowledge points will be directly reflected in the graph and play a crucial role in the subsequent work. In this study, the Word2vec algorithm was used to extract words with strong semantic relations, the association rule algorithm was used as a supplement and, finally, the string suffix matching method was used to extract categorical relations.

In this study, we used Word2vec to obtain entity relationships by semantic computation of text content, and the larger the semantic similarity value between words, the closer the relationship between words. Before training the model, we needed to pre-process the corpus, which included removing useless characters from the data, dividing the data into words, removing deactivated words and noisy words, filtering the word nature, etc.; we also stitched all of the documents to form a large document and then used the Skip-Gram model to train the documents with word vectors; we finally converted the document files into word vector files. The second step was to use the obtained entity set to calculate the cosine similarity between each entity and other entities, save the entity relationship pairs that were larger than the relevant cosine similarity threshold and, finally, judge the candidate entity relationships and save the ones that meet the requirements to form the candidate entity relationship set.

Text processing is required before using the association rule algorithm, using the obtained entity set for text matching, if a word is displayed once, this word is recorded, and after that, it appears again without increasing only once, and multiple records are obtained; each record shows which domain entities appear in each document, and then all record numbers are obtained to form the set of transactions to be used. The reasonable settings of support and confidence can effectively get strong association rules, i.e., entity relationship pairs. The values of support and confidence depend on the size of the dataset. If the support and confidence are set too high, the entity pairs that originally have related relationships will be screened out, and if they are set too low, the wrong entity relationship pairs will be included, thus reducing the accuracy of relationship extraction; so, it is necessary to set reasonable support and confidence through several experiments and finally determine all of the association with low support and confidence rule pairs, as shown in Figure 8.

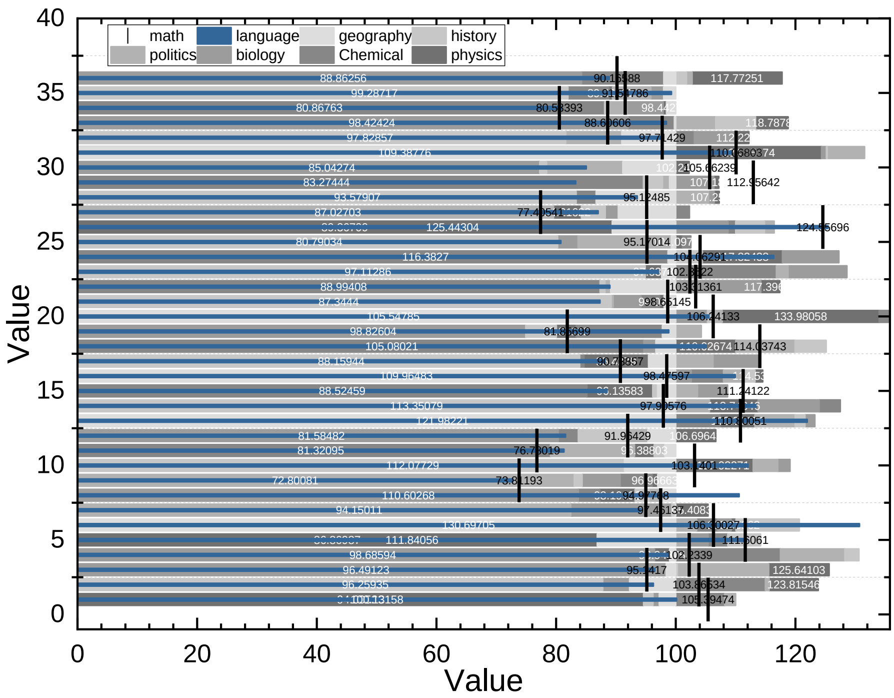

Figure 8. Accuracy of the entity.

In terms of the overall results, the accuracy of the overall experiment was above 80%, and even exceeded 90% in some subjects. From the overall accuracy, the entity extraction method used in this study is extremely helpful in improving the accuracy of entities in the field of basic education. The extraction results will be different among different subjects, but the differences were not big here; among them, the best accuracy was observed for political science and the worst was obtained for physics. The difference value of accuracy between two subjects reached 10%, and the effects of language, mathematics, chemistry, biology, politics and geography were better than those for physics and history. Compared with the previous group test, the final physical accuracy of the group for language and physics subjects was 68.59 and 85.08%, respectively. Obtaining valuable information by means of knowledge graph visualization can be studied, but there are relatively few studies on the application of knowledge graphs for information display in colleges and universities. Compared with this study, the accuracy of language subjects has improved greatly and that of physics subjects has decreased, but not by much.

To verify the validity of the knowledge map, the corresponding evaluation indexes in the field of basic education were constructed according to other evaluation indexes, and a comparative analysis of the eight disciplines in which a knowledge map for the field of basic education was conducted. The specific evaluation dimensions and indicators have been introduced, and the experimental results of each discipline were analyzed according to the corresponding indicators. Finally, through experimental verification, the overall effect of this method was found to be better in the usability dimension, and the construction method of this work showed a good improvement in accuracy; but, for the other indicators, the degree of improvement was found to be very different in different disciplines. The construction result of the language discipline was better than that of the physics discipline, and the performance of the construction result of the knowledge map of this paper was different for different indicators from different disciplines.

5. Conclusions

As a good representation of knowledge structure, knowledge graphs have been widely used in the field of basic education. The research direction of this study has been proposed for the current application problems of knowledge graphs in the field of basic education and the quality problems corresponding to the knowledge graphs in the field of basic education. Concerning the visualization model, functional requirements have been proposed and functional modules were divided. Strictly according to the implementation scheme of each module, the relevant system framework was formulated; the results of the visualization model were realized and displayed by combining the existing technologies. In the process of entity extraction using the above method, it was found that the threshold value setting was not the same for different disciplines, and only when the threshold value is set correctly will the overall extraction effect be more helpful. In addition, an increase in the number of corpora will improve the accuracy of entity extraction results to a certain extent. The visualization model design is based on the knowledge graph of university information. By designing the model in consideration of two aspects, i.e., reducing the complexity of graphical element structure and user interaction, a more detailed user interest model, keyword search model and interaction model were designed and used as the reference for designing the knowledge graph visualization platform for college information according to the feasibility. The construction method of the knowledge graph was comprehensively analyzed and applied to the education field. The basic information knowledge graph of colleges and universities was constructed by using relevant construction techniques. Using the knowledge map visualization of college information as an application scenario, the nodes in the layout were reduced by cropping, and the visualization model was designed to provide solution ideas for solving the problems of visual confusion, knowledge disorientation and inability to focus on user interests that exist in large-scale knowledge map visualization.

Acknowledgments

This work was supported by the Foundation of Humanities and Social Sciences of Ministry of Education of China (No. 19YJC603114); the Education Research Project of Philosophy and Social Science in Yunnan Province (No. AD21001) and the Science Research Fund Foundation of Yunnan University of Finance and Economics (No. 2020C03; 2021C35).

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. Z. Guo, K. Yu, Z. Lv, K. K. R. Choo, P. Shi, J. J. P. C. Rodrigues, Deep federated learning enhanced secure POI microservices for cyber-physical systems, *IEEE Wireless Commun.*, **29** (2022), 22−29. https://doi.org/10.1109/MWC.002.2100272
2. Y. He, S. Yang, C. Y. Chan, L. Chen, C. Wu, Visualization analysis of intelligent vehicles research field based on mapping knowledge domain, *IEEE Trans. Intell. Transp. Syst.*, **22** (2020), 5721−5736. https://doi.org/10.1109/TITS.2020.2991642

3. L. Zhao, Z. Yin, K. Yu, X. Tang, L. Xu, Z. Guo, et al., A fuzzy logic based intelligent multi-attribute routing scheme for two-layered SDVNs, *IEEE Trans. Netw. Serv. Manage.*, **2022** (2022). https://doi.org/10.1109/TNSM.2022.3202741
4. Y. Li, H. Ma, L. Wang, S. Mao, G. Wang, Optimized content caching and user association for edge computing in densely deployed heterogeneous networks, *IEEE Trans. Mob. Comput.*, **21** (2022), 2130−2142. https://doi.org/10.1109/TMC.2020.3033563
5. C. Chen, Z. Liao, Y. Ju, C. He, K. Yu, S. Wan, Hierarchical domain-based multi-controller deployment strategy in SDN-enabled space-air-ground integrated network, *IEEE Trans. Aerosp. Electron. Syst.*, **58** (2022), 4864−4879. https://doi.org/10.1109/TAES.2022.3199191
6. Z. Zhou, X. Dong, Z. Li, K. Yu, C. Ding, Y. Yang, Spatio-temporal feature encoding for traffic accident detection in VANET environment, *IEEE Trans. Intell. Transp. Syst.*, **2022** (2022), 19772−19781. https://doi.org/10.1109/TITS.2022.3147826
7. Z. Guo, C. Tang, W. Niu, Y. Fu, T. Wu, H. Xia, et al., Fine-grained recommendation mechanism to curb astroturfing in crowdsourcing systems, *IEEE Access*, **5** (2017), 15529−15541. https://doi.org/10.1109/ACCESS.2017.2731360
8. S. Xia, Z. Yao, Y. Li, S. Mao, Online distributed offloading and computing resource management with energy harvesting for heterogeneous MEC-enabled IoT, *IEEE Trans. Wireless Commun.*, **20** (2021), 6743−6757. https://doi.org/10.1109/TWC.2021.3076201
9. Q. Zhang, K. Yu, Z. Guo, S. Garg, J. J. P. C. Rodrigues, M. M. Hassan, et al., Graph neural networks-driven traffic forecasting for connected internet of vehicles, *IEEE Trans. Network Sci. Eng.*, **9** (2022), 3015−3027. https://doi.org/10.1109/TNSE.2021.3126830
10. S. Ji, S. Pan, E. Cambria, P. Marttinen, P. S. Yu, A survey on knowledge graphs: representation, acquisition, and applications, *IEEE Trans. Neural Networks Learn. Syst.*, **33** (2021), 494−514. https://doi.org/10.1109/TNNLS.2021.3070843
11. Y. Yang, Y. Zhuang, Y. Pan, Multiple knowledge representation for big data artificial intelligence: framework, applications, and case studies, *Front. Inf. Technol. Electron. Eng.*, **22** (2021), 1551−1558. https://doi.org/10.1631/FITEE.2100463
12. W. Wang, C. Lu, Visualization analysis of big data research based on Citespace, *Soft Comput.*, **24** (2020), 8173−8186. https://doi.org/10.1007/s00500-019-04384-7
13. M. Decuypere, P. Landri, Governing by visual shapes: university rankings, digital education platforms and cosmologies of higher education, *Crit. Stud. Educ.*, **62** (2021), 17−33. https://doi.org/10.1080/17508487.2020.1720760
14. F. Lecue, On the role of knowledge graphs in explainable AI, *Semantic Web*, **11** (2020), 41−51. https://doi.org/10.3233/SW-190374
15. R. Cerezo, A. Bogarín, M. Esteban, C. Romero, Process mining for self-regulated learning assessment in e-learning, *J. Comput. High Educ.*, **32** (2020), 74−88. https://doi.org/10.1007/s12528-019-09225-y
16. L. Chang, G. Zhou, O. Soufan, J. Xia, miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology, *Nucleic Acids Res.*, **48** (2020), W244−W251. https://doi.org/10.1093/nar/gkaa467
17. H. Luan, C. C. Tsai, A review of using machine learning approaches for precision education, *Educ. Technol. Soc.*, **24** (2021), 250−266. Available from: https://www.jstor.org/stable/26977871.
18. D. Ifenthaler, J. Y. K. Yau, Utilising learning analytics to support study success in higher education: a systematic review, *Educ. Technol. Res. Dev.*, **68** (2020), 1961−1990. https://doi.org/10.1007/s11423-020-09788-z

19. J. Yuan, C. Chen, W. Yang, M. Liu, J. Xia, S. Liu, A survey of visual analytics techniques for machine learning, *Comput. Visual Media*, **7** (2021), 3−36. https://doi.org/10.1007/s41095-020-0191-7
20. N. Pellas, S. Mystakidis, I. Kazanidis, Immersive virtual reality in K-12 and higher education: a systematic review of the last decade scientific literature, *Virtual Reality*, **25** (2021), 835−861. https://doi.org/10.1007/s10055-020-00489-9
21. P. Jääskelä, V. Heilala, T. Kärkkäinen, P. Häkkinen, Student agency analytics: learning analytics as a tool for analysing student agency in higher education, *Behav. Inf. Technol.*, **40** (2021), 790−808. https://doi.org/10.1080/0144929X.2020.1725130
22. L. Tarkhova, S. Tarkhov, M. Nafikov, I. Akhmetyanov, D. Gusev, R. Akhmarov, Infographics and their application in the educational process, *Int. J. Emerging Technol. Learn. (IJET)*, **15** (2020), 63−80. https://doi.org/10.3991/ijet.v15i13.14647
23. H. Gao, Y. Duan, L. Shao, X. Sun, Transformation-based processing of typed resources for multimedia sources in the IoT environment, *Wireless Netw.*, **27** (2021), 3377−3393. https://doi.org/10.1007/s11276-019-02200-6
24. M. Koho, E. Ikkala, P. Leskinen, T. Tamper, J. Tuominen, E. Hyvönen, Warsampo knowledge graph: Finland in the second world war as linked open data, *Semantic Web*, **12** (2021), 265−278. https://doi.org/10.3233/SW-200392
25. I. Krak, O. Barmak, E. Manziuk, Using visual analytics to develop human and machine-centric models: a review of approaches and proposed information technology, *Comput. Intell.*, **38** (2022), 921−946. https://doi.org/10.1111/coin.12289