Dynamic Knowledge Graph: A Tool for Fostering Conceptual Change in Knowledge Building Community

Zengcan Xue, Nanjing Normal University, Nanjing
Yibing Zhang, Nanjing Normal University, Nanjing
Chaojun Xu, Nanjing Normal University, Nanjing
Email: xue_zc@126.com, zhyb304@126.com, 267440207@qq.com

Abstract: In this short paper, we introduce a work-in-process tool called Dynamic Knowledge Graph (DKG), which is designed to facilitate conceptual change in knowledge building environments. The main function of DKG is to support higher conceptual change by turning current artifacts in the form of Knowledge Forum notes into instructional resources. DKG can provide relevant dynamic knowledge graphs that reflect core concepts and its relationships that are mined from current artifacts for learners. Preliminary study suggested that the DKG could be further improved and we conclude this paper by discussing its current limitations and future directions.

Introduction

Knowledge building (KB) is defined as the production and continual improvement of ideas of value to a community (Scardamalia & Bereiter, 2014). It attaches importance to conceptual engagement and contribution, which can help students to obtain the essence of scientific concept and develop new knowledge. Efforts are made in the KB community to advance conceptual change. Burtis, Chan, Hewitt, Scardamalia and Bereiter (1993) investigated this possibility that KB fosters conceptual change of students using CSILE, a predecessor to Knowledge Forum (KF). Furthermore, students who maintained knowledge-building goals and evaluated their ideas in the context of the writings of the scientific community beyond their classroom succeeded in attaining critical conceptual change (Oshima & Scardamalia, 1996). Especially, high-conceptual-progress students in KB were more concerned with constructing their knowledge centered around problems, whereas low-conceptual-progress students were more involved in accumulating referent-based knowledge (Oshima, Scardamalia & Bereiter, 1996).

On the other hand, KB as a mediator of conflict in conceptual change helps to highlights the importance of students' constructive activity in learning (Chan, Burtis & Bereiter, 1997). In one study, when students inquire and reflect on their understanding in the context of KB community knowledge, they may advance in their individual and group learning that predicted subsequent conceptual change (Lam & Chan, 2008). Likewise, students in KB were able to recognize a gap or conflict in their knowledge and willingly sought information to improve their naive conceptions (Khanlari, Bereiter & Scardamalia, 2016). A further educationally significant finding is the willingness of students to take collective cognitive responsibility to improve the knowledge of the community and solve their peers' misconceptions. If KB technology is indeed becoming smarter, supports for self-organization around conceptual contributions to encourage emergent conceptual change are needed. What is missing in current knowledge-building environments is a mechanism to explicitly support such collaborate conceptual change using data-intensive analytics (Chen & Liu, 2016). To this end, we developed DKG in knowledge building community to explore this possibility.

Knowledge Graph in Conceptual Change

As an integrated information repository, knowledge graph interconnect heterogeneous data from different domains. A prominent example is Google's knowledge graph, which represents real-world entities and relationships through multiple relational graphs (Chen et al., 2018; Rizun, 2019). In education domain, knowledge graphs are often used in school subject teaching, also known as concept maps. For students, understanding and applying the logical relationships between entities or concepts requires more cognitive engagements. Because the nodes of the knowledge graph link entity information, it can promote deeper learning within a certain domain. Moreover, due to the concealment of group information and learning processes, it is necessary to visualize the consensus content through group awareness, so as to promote active learning (Bodemer & Dehler, 2011). Using the knowledge graph, teachers and students can interact with others through knowledge nodes and relations, explore collaborations and track rea-time changes (Rafols, Porter, & Leydesdorff, 2010).

Several studies have clearly demonstrated the effectiveness of knowledge graphs in supporting the conceptual change. Knowledge graphs can reveal interesting information about the process of knowledge restructuring that is assumed to happen in conceptual change, which have been used in science education as tools for supporting students' learning of the structural nature of physics knowledge but also as tools for assessment and evaluation of learning (Goldwater & Schalk, 2016). The method of knowledge graphs emerged to leverage the understanding of the process of conceptual change in science (Novak & Musonda, 1991), growing out of theories describing cognitive structures recognizing the interrelatedness of concepts as an essential property of knowledge (Ruiz-Primo & Shavelson, 1996). When knowledge graphs are used repeatedly over time, conclusions concerning the development of knowledge can be drawn. How knowledge develops for a group of learners can reveal information about core concepts and misconceptions that may hinder the learning progress at a certain stage and which teachers can then approach in the classroom. Moreover, knowledge graphs can yield valuable information for teachers by allowing them to see common elements that their students did not yet fully understand or that are prone to misconceptions (Duit & Treagust, 2003). In a longitudinal study, the changes in students' knowledge structures were examined through knowledge graphs. Besides a growth of the knowledge network, the results indicated a reorganization, with first a fragmentation during the unit, followed by an integration of knowledge at the end of the unit. Moreover, the terms used in the knowledge graphs varied in their centrality, with more abstract terms being more central and thus more important for the structure of the graph (Thurn, Hänger & Kokkonen, 2020).

Such knowledge graphs are usually constructed by experienced teachers or domain experts in manual way. However, such a manual construction process is actually time-consuming and not scalable to large number of concepts and relations. On the other hand, the manual construction approach is error-prone: according to the pedagogical research, there often exists expert blind spot (Nathan & Petrosino, 2003), which means expert's cognition and learner's cognition on the same concept often do not well align. As a result, those manually created knowledge graphs can be suboptimal or misleading for learners.

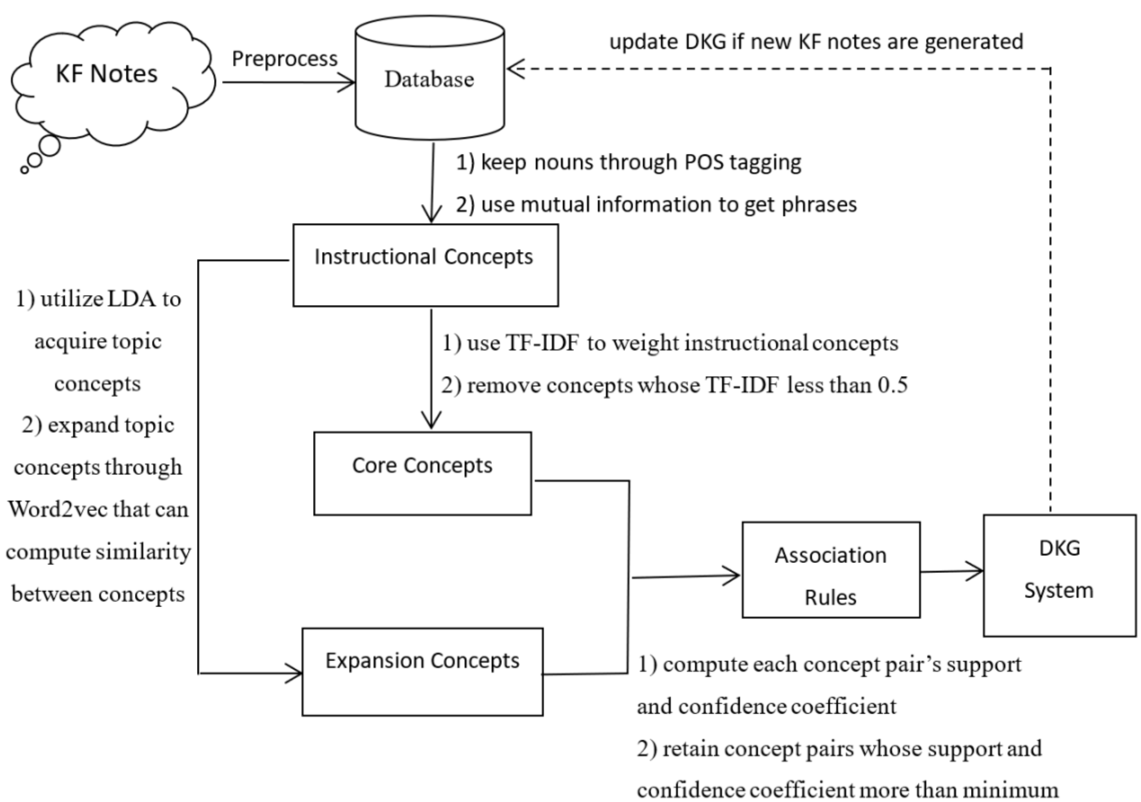

Figure 1. A flow chart of the DKG.

Dynamic Knowledge Graphs

Motivated by the increasing demands for knowledge graph in KF and the limitations of manual construction approach, we propose DKG to automatically construct educational dynamic knowledge graphs that can be used for teaching and learning in knowledge building community. To begin with, the desired nodes in knowledge graphs of KF represent instructional concepts in subjects or courses, so the extraction requires KF data from students' notes. In addition, the relations between instructional concepts reflect learner's cognitive and educational process. Such relations are relatively difficult to identify without proper analysis and modeling on the specific KF data. Below, we briefly explain the current technical implementation of DKG.

The algorithmic computation that powers DKG is illustrated in Figure 1. The computation includes three phases. First, the main function of primary stage is to extract instructional concepts from KF notes, which need to be firstly converted from KF notes into machine-readable text format. A part-of-speech (POS) tagging algorithm can be employed to extract particular concepts. And then TF-IDF (Salton & Buckley,1988), LDA (Blei, Ng & Jordan, 2003) and Word2vec (Li et al., 2019) are used to select core concepts and expand the size of concepts corpus. To be specific, TF-IDF algorithm can sort high-frequency core concepts, afterwards we retain the concepts whose TF-IDF value is higher than 0.5. In addition, LDA can find the subject terms from the semantic perspective, which can be expanded through Word2vec deep learning package.

The second phase is to identify the educational relations that interlink instructional concepts to help the knowledge building process directly. Since educational relations are abstract, this module utilizes data mining technique, such as the association rule mining (Agrawal & Srikant, 1994). The number of texts appearing in A word and B word is C_A and C_B respectively, the total number of texts is C_T. Then, according to the co-occurrence matrix, the total number of $C_{A\cap B}$ occurring in all texts for each word pair {A and B} is obtained. The support coefficient describes the words A and B appear simultaneously in all texts is calculated as follows: Support $(A\rightarrow B) = P(A\cap B) = C_{A\cap B}/C_T$. The confidence coefficient describes the probability of the occurrence of the text of word A and word B. The calculation formula is as follows: Confidence $(A\rightarrow B) = P(B|A) = C_{A\cap B}/C_A$. For each word pair $\{A\rightarrow B\}$ in this study, the support is greater than or equal to the minimum support, and the confidence is greater than or equal to the minimum confidence.

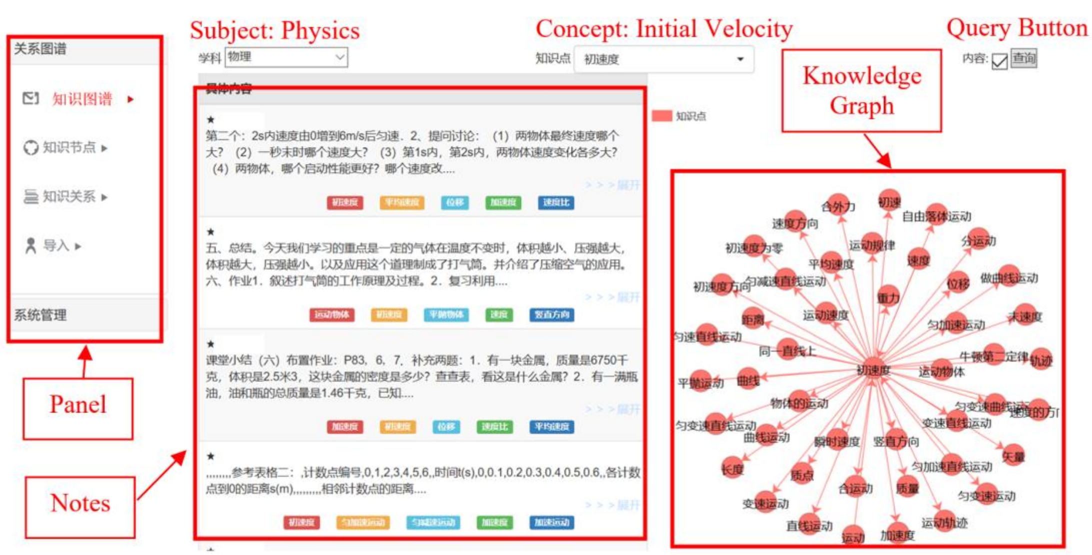

Figure 2. A snapshot of the DKG system. The panel includes four functions: knowledge graphs, knowledge nodes, knowledge relationships and import function. In the knowledge graph page, the user can select the specific instructional concept (e.g., initial velocity), and then click the query button to view the relevant knowledge graph and notes, which can help users track changes in the structure of core concepts.

Finally, dynamic knowledge graphs are displayed in the knowledge graph management system (see Figure 2 and Figure 3). In Figure 2, the core concepts and its relations returned by the computation. In this display, the users can simply select a concept to query related knowledge graphs and corresponding note contents, meanwhile, the users can manually add knowledge nodes and establish knowledge relationship, which is a thoughtful design. In Figure 3, the users have access to the whole knowledge graph of one certain subject in KF, who will have a clear idea of the position and importance of a particular concept in the overall knowledge graph, as well as its relations with other concepts.

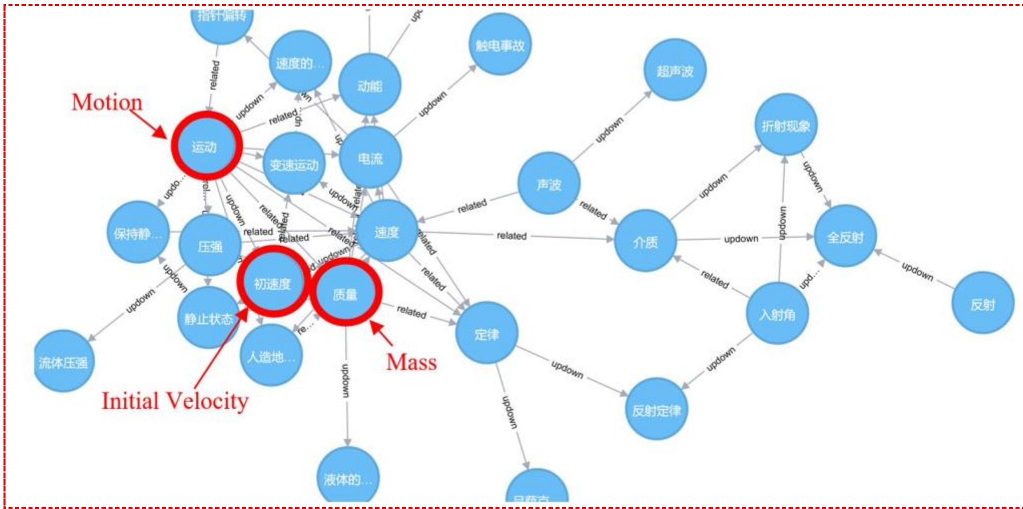

Figure 3. Exemplary Knowledge Graph of Physics in DKG. This knowledge graph is constructed from the KF notes of the movement chapter in junior middle school physics. We can find that the instructional concepts of motion, mass and initial velocity are all at the core of the knowledge graph, which are of great significance for fostering conceptual change in knowledge building.

Preliminary Studies

A knowledge graph can be regarded as a network consisting of nodes (also called vertices) and links (also called edges) between them, which enables the application of graph theory on knowledge graphs (Chen, Chang, Ouyang & Zhou, 2018). Regarding the centrality of certain nodes, several measures exist, from counting the number of edges per node (degree centrality), over the average shortest path length to all other nodes (closeness centrality) to the number of paths that cross through a certain node (betweenness centrality). Each centrality measure answers a slightly different question. Moreover, as a comparison measure, we applied PageRank centrality that takes the number of (incoming) edges but also the importance of adjacent nodes sending these edges into account. In contrast to the aforementioned centrality indices, PageRank centrality weights each edge differently according to the importance the node it emerges from.

We used Python for all analyses, with the following packages in alphabetical order: igraph, json, lxml, matplotlib, numpy, pylab, re, scipy, scrapy, xlwt. As centrality indices strive to identify the most influential nodes, they are less suited for the nodes with low centrality and do not necessarily express a meaningful order of such nodes with lower centrality because of a lack of sensitivity. Thus, we focus on the centrality statistics for the representative nodes, such as motion, mass and initial velocity (See Figure 3). The result of graph analysis is reflected in Table 1.

Table 1: Graph analysis: Most central nodes indicated by degree, closeness, betweenness and PageRank given for comparison.

Central Nodes	Degree	Closeness	Betweenness	PageRank
Motion	21	0.0124	1028	0.049
Mass	8	0.0123	959	0.022
Initial Velocity	2	0.0123	1.537	0.004

The term conceptual change is often associated with changes in the deeper, underlying knowledge. The process of conceptual change, however, comes in degrees. At the most basic level, conceptual change is associated with assimilation of new knowledge and facts into the existing knowledge structure, enriching it (Vosniadou, 1994). In terms of Table 1, the concept of motion has the largest degree, indicating that it occupies an important position in the network. By contrast, the degree of mass ranks second, and the degree of initial velocity is the smallest, indicating that initial velocity is less important than motion and mass, which is also consistent with the topic of "movement in physics". Furthermore, the closeness of the three concepts imply that they are all relatively close to the other nodes. As for betweenness and PageRank, motion and mass are similar, but the value of initial velocity is smaller. There is a proof that explains strong connections between motion or mass and other instructional concepts. However, initial velocity shows less potential to link other important concepts. According to Newton's first law (the law of inertia), the relationship between motion and mass is very tight, which are also crucial nodes in the movement chapter in junior middle school physics. It is natural to conclude that the graph analysis results conform to the corresponding instructional content and can reveal the core concepts structure and its relations, which input impetus to help KB teachers track and compare the changes of students' conceptual networks over time. In other words, the dynamic knowledge graph shows great prospect for promoting and evaluating students' conceptual change in the knowledge-building classrooms.

Discussion and Conclusions

In this short paper, we introduce DKG, an ongoing feature, developed to generate knowledge graphs in a knowledge-building community. The central affordance of DKG is to turn KF notes in a community into resources for continuous knowledge building, by feeding pertinent knowledge graphs to learners. The algorithm that powers DKG is not complex and open to further refinement. But the tool itself is of significance in facilitating conceptual change with dynamic knowledge graphs in knowledge building. Benefits of such scaffolding mechanisms are demonstrated in earlier design research using networks (Oshima J, Oshima R, & Matsuzawa, 2012). With DKG, we attempt to make such scaffolding efforts even more dynamic, concurrent and automatic. Planning of new design research initiatives is underway to develop pedagogical principles for incorporating DKG in knowledge-building classrooms.

In the first step, teachers can examine what concepts are unknown before teaching. To survey to what extent new concepts are understood, teachers could examine whether new concepts are meaningfully incorporated into the knowledge graph and check on possible misconceptions. Similarly, if the centrality analysis is adopted, the important concepts of high centrality can also be considered to diagnose the change of students' knowledge structure. Teachers can also ask students to compare the final knowledge graph with the expert knowledge graph. Students can be prompted to identify important differences between the two graphs and to reflect on what they can improve. Such comparison can enable students to deeply analyze their own ideas from the perspective of experts, thus helping them to better integrate new knowledge.

In addition to graph analysis reported earlier, there are several challenges we need to figure out. First, the integration and real-time performance of DKG need to be further improved. Currently, DKG relies on manual import of KF notes, so there is no way to seamlessly carry out the whole process of data collection, graph construction, graph analysis and graph update. Second, association rules algorithm should be further subdivided. According to the degree of support and confidence coefficient, association rules can identify three types of relationships: basic relationship, advanced relationship and peer relationship. Finally, we want to give users, including teachers and students, a chance to bring an impact on DKG. By doing so, future versions of DKG will solicit input from users, thus "collaborating" with them rather than forcing advice on them. Besides, students' construction of dynamic knowledge graphs through DKG can potentially provide important motivation for long-term knowledge building and higher level of conceptual change.

References

Agrawal R, Srikant R. (1994). Fast Algorithms for Mining Association Rules. *Proceedings of the 20th International Conference on Very Large Data Bases* (pp. 487-499). Santiago, Chile: IEEE.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(1), 993-1022.

Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, *27*(3), 1043–1045.

Burtis, J., Chan, C., Hewitt, J., Scardamalia, M., & Bereiter, C. (1993). Misconceptions in a student-guided research project. Poster presented at *the annual meeting of the American Educational Research Association.* Atlanta: AERA.

Chan C, Burtis J, Bereiter C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and instruction*, *15*(1): 1-40.

Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education*, *37*, 21-30.

Chen B, Liu X. (2016). WordWhispers: A tool for bootstrapping vocabulary development in knowledge-building communities. Paper Present at *the Knowledge Building Summer Institute*, Singapore: Institute for Knowledge Innovation and Technology.

Chen, P., Lu, Y., Zheng, V. W., Chen, X., & Yang, B. (2018). Knowedu: A system to construct knowledge graph for education. Ieee Access, 6: 31553-31563.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International journal of science education*, *25*(6), 671-688.

Goldwater M B , Schalk L. (2016). Relational Categories as a Bridge Between Cognitive and Educational Research. *Psychological Bulletin*, *142*(7):729-757.

Khanlari A, Bereiter C, Scardamalia M. (2016). Misconceptions and conceptual change in knowledge building environments. Paper Present at *the Knowledge Building Summer Institute*, Singapore: Institute for Knowledge Innovation and Technology.

Lam, I. C. K., & Chan, C. K. K. (2008). Fostering epistemological beliefs and conceptual change in chemistry using knowledge building. *Proceedings of the 8th international conference on international conference for the learning sciences* (pp. 461-468). Utrecht, The Netherlands: International Society of the Learning Sciences.

Li, B., Drozd, A., Guo, Y., Liu, T., Matsuoka, S., & Du, X. (2019). Scaling word2vec on big corpus. *Data Science and Engineering*, *4*(2), 157-175.

Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal*, *40*(4), 905-928.

Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American educational research journal*, *28*(1), 117-153.

Oshima J, Oshima R, Matsuzawa Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse[J]. *Educational Technology Research and Development*, *60*(5): 903-921.

Oshima J, Scardamalia M, Bereiter C. (1996). Collaborative learning processes associated with high and low conceptual progress. *Instructional Science*, *24*(2): 125-155.

Oshima, J., & Scardamalia, M. (1996). Knowledge-building and conceptual change: an inquiry into student-directed construction of scientific explanations. In *Proceedings of the 1996 international conference on Learning sciences* (pp. 256-263). Illinois: International Society of the Learning Sciences.

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, *61*(9), 1871–1887.

Rizun M. (2019). Knowledge graph application in education: a literature review. *Acta Universitatis Lodziensis. Folia Oeconomica*, *3*(342): 7-19.

Ruiz-Primo M A, Shavelson R J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, *33*(6): 569-600.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management*, *24*(5), 513-523.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (2nd ed., pp. 397–417). New York, NY: Cambridge University Press.

Thurn, C. M., Hänger, B., & Kokkonen, T. (2020). Concept Mapping in Magnetism and Electrostatics: Core Concepts and Development over Time. *Education Sciences*, *10*(5), 129.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and instruction*, *4*(1), 45-69.

Acknowledgments

Support for research and writing this article was provided by the China National Social Science Foundation: Research on the Construction Method and Application Model of Knowledge Graph in Basic Education (Grant Number: BCA190087).