Construction of Curriculum Knowledge Graph for Educational Informatization

Guoming Liao[1,*], Ji Chen[2,a]

[1]College of digital Engineering, Chongqing College of Architecture and Technology, Chongqing, China
[2]Research and Development Department, Chongqing Hanhai Wisdom Big Data Technology Co. LTD, Chongqing, China

*1643493318@qq.com, a44540155@qq.com

Abstract. In response to issues such as incomplete design of knowledge points, scattered knowledge points, and single relationship of knowledge points in professional courses, a course knowledge system based on knowledge graph technology is proposed. Taking Hadoop courses as an example, based on the Neo4j graph database, a knowledge graph of Hadoop courses is constructed from four stages: course knowledge point extraction, course knowledge point relationship extraction, course knowledge structure organization, and course knowledge point attribute addition. The visualization of the course knowledge graph is carried out, and then the constructed knowledge graph is applied to course knowledge learning of students, learning path recommendation, and other aspects to achieve personalized teaching.

Keywords: knowledge graph, Hadoop course, Neo4j graph database, personalized teaching

1 Introduction

The development history of knowledge graph can be traced back to the 1960s. At that time, one of the pioneers of artificial intelligence, John Sowa, proposed the concept of "semantic network". He believed that knowledge could be represented and organized through the correlation between nodes and edges. With the development of computer technology, knowledge graph has been more widely applied and developed.

With the rapid development of educational informatization, the application of knowledge graph technology in the field of education has become a current development trend. The knowledge graph has played an important role in personalized learning, intelligent educational assistants, personalized generation of teaching content, teaching evaluation and feedback, and teaching improvement and optimization. It will further promote the development and innovation of educational technology.

Y. Kuang et al. (eds.), *Proceedings of the 2024 5th International Conference on Education, Knowledge and Information Management (ICEKIM 2024)*, Atlantis Highlights in Computer Sciences 22,
https://doi.org/10.2991/978-94-6463-502-7_53

2 Knowledge Graph Technology

In May 2012, the term "knowledge graph" was officially proposed by Google. It uses entities, entity attributes, and entity relationships to describe various conceptual entities in the objective world and their relationships. It is widely used in knowledge organization and semantic modeling research [1]. The original purpose of knowledge graph proposed was to optimize search engines, but with rapid development in recent years, it has been widely applied in fields such as education, e-commerce, etc. It plays a vital role in intelligent search, intelligent question answering, intelligent decision-making, and recommendation systems [2-3].

The knowledge graph is a newly emerging discipline and research area in the field of artificial intelligence. Building a course knowledge graph is an important research content in the fields of knowledge graph, network learning, and knowledge services. Its foundation is a knowledge base of relational data, marking related data to clarify relationships, and creating a low-level knowledge structure network [4]. The main technologies of knowledge graph are knowledge representation, knowledge extraction, knowledge fusion, knowledge reasoning and graph application [5]. The knowledge extraction, also known as information extraction, is an automatic extraction technique that extracts entities, relationships, attributes, and other data from semi-structured and unstructured data [6]. After knowledge extraction, data such as entities, their relationships, and attributes are obtained, but the logic, completeness, and accuracy of this information and data are not perfect, and it still needs to be integrated through knowledge fusion [7]. The new course resource organization form based on knowledge graph can associate scattered knowledge points with scattered teaching resources and knowledge points, integrate multiple modes of teaching resources, and provide advanced teaching technology means such as user portrait, personalized recommendation and decision support[8-10].

3 Construction of Hadoop Course Knowledge Graph

3.1 Extracting Knowledge Points from Hadoop Courses

When determining knowledge points, this paper strives to be as detailed and specific as possible, avoiding knowledge points being too broad or repetitive. At the same time, it also considers the logical relationships and hierarchical structure between knowledge points, so that the relationships between knowledge points can be clearly represented when constructing a knowledge graph in the future.The core knowledge points of the course are determined by carefully consulting teaching materials, course standards and other teaching resources.The extracted knowledge points are shown in Table 1.

Table 1. Hadoop Course Knowledge Extraction

Course Knowledge Module	Knowledge Points of Partial Courses
Big Data with Hadoop	Big data characteristics, relationship with Hadoop, etc
Hadoop Deployment	CDH, Linux, Vmware, Centos, IP, Host Name, etc
Hadoop Components	Distributed File System, NameNode, DataNode, etc
Zookeeper Distributed Collaboration Framework	Zookeeper Service, Zookeeper Role, etc
Hadoop HA Theory and Deployment	Hadoop HA Introduction, HA Cluster Planning,etc

3.2 Extracting the Relationship Between Knowledge Points in Hadoop Courses

When determining the knowledge relationship of Hadoop courses, this paper refers to teaching resources such as textbooks, course outlines, teacher handouts, and opinions from experts in the field of big data. At the same time, it also draws on the experience of the knowledge relationships in other relevant courses and teaching standards to ensure the accuracy and completeness of the knowledge graph. This article mainly extracts the knowledge point relationships of Hadoop course from the following aspects.

One is to ascertain the relationship between parent and child nodes based on the hierarchical relationship between knowledge points. The second is to determine the correlation relationship based on the correlation between knowledge points. The third is to determine the order based on the dependency relationship between knowledge points. The fourth is to determine similarity relationships based on the similarity between knowledge points.

3.3 Organization of Hadoop Course Knowledge Structure

When organizing the knowledge structure of Hadoop courses, the Neo4j graph database is used to manage and display the knowledge structure in the knowledge graph.

One is to divide the course content into several modules, each containing a set of related knowledge points.The second is to determine the hierarchical structure of the knowledge structure based on the hierarchical relationship between knowledge points.The third is to make sure the order of knowledge structure based on the dependency relationship and learning order between knowledge points.The fourth is to learn certain knowledge points in the course in parallel without obvious order.The fifth is to determine key knowledge points based on the course objectives and learning requirements. These key knowledge points can be the core concepts, key technologies, or common problems of the course. This paper divides the key knowledge points into a separate course module for setting.

3.4 Adding Attribute Information of Hadoop Course Knowledge Points

When constructing the Hadoop course knowledge graph, it is necessary to add attribute information such as name, description, difficulty level, examples, and applications to each knowledge point to provide more context and detailed information. Taking the knowledge point - distributed file system as an example, the relevant attributes in Table 2 should be added to it.

Table 2. Knowledge Points of Hadoop Course - Attributes of Distributed File Systems

Attribute	Attribute Information
Name	Distributed File System
Description	A distributed file system is a system that stores files and data on multiple computer nodes.
Examples and Applications	HDFS in Hadoop is a typical distributed file system
Related Resources	Training Guide Manuals, Documents, and other Resources
Learning Path	Big data characteristics→ Basic concepts of Hadoop → Distributed file systems → Hadoop HDFS
Difficulty Level	Medium

4 Visualization of Hadoop Course Knowledge Graph

This paper utilizes the Neo4j graph database for storing Hadoop course knowledge points. Unlike traditional relational databases that store data in database table fields, graph databases store relationships between data and data in nodes and edges, which are referred to as "nodes" and "relationships" in graph databases. Any relationship consists of a start node, an end node, and an edge that points from the start node to the end node. All nodes in the database are connected through various relationships, and the final Hadoop course knowledge graph visualization is shown in Figure 1, while the Hadoop course sub module knowledge graph visualization is shown in Figure 2.

Fig. 1. Partial Screenshot of Hadoop Course Knowledge Graph

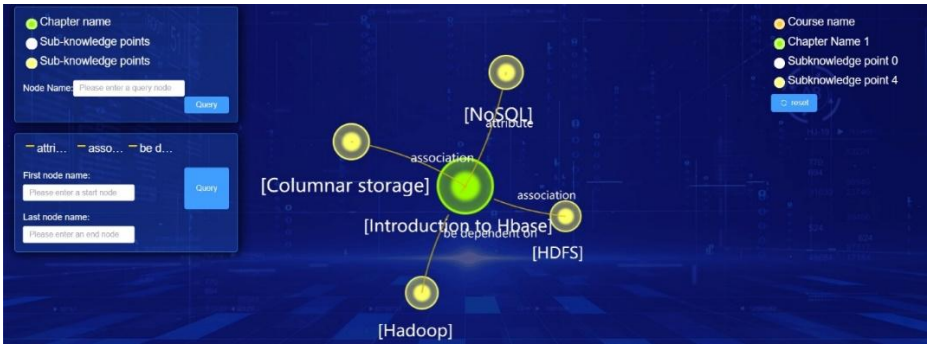

Fig. 2. A Screenshot of the Knowledge Graph for Some Sub Modules of Hadoop Course

5 Update and Maintenance of Knowledge Graph

Updating and maintaining the course knowledge graph is a continuous process, which follows a series of processes including collection of feedbacks, analysis and evaluation, update of knowledge points, update of attribute information, optimization of association relationships, update of resource links, regular maintenance and continuous improvement.Through continuous updating and maintenance, the timeliness, accuracy and applicability of the curriculum knowledge graph can keep up with the times and better meet the needs and teaching objectives of learners.

6 Conclusion

This paper provides a detailed introduction to the construction process of Hadoop course knowledge graph, mainly including knowledge extraction, knowledge point relationship extraction, knowledge structure organization, knowledge point attribute addition, and knowledge graph visualization. By accurately extracting the content of

textbooks, syllabuses and network resources in the course, and promoting the visualization of subject knowledge with the help of pictures and other forms, the combination of multiple knowledge points can reflect the structure and distribution of knowledge points, which has positive significance for students to build a clear course knowledge system, and also helps to enhance the interest of knowledge learning. Knowledge graph is the core driving force to promote the development of artificial intelligence. By extending the knowledge graph studied in this paper, intelligent search, knowledge inference and learning path recommendation can be realized through the construction of subject knowledge graph, which will provide new empowering power for education and teaching in the era of education informatization 2.0.

References

1. Wan HP, Cheng LN, Cheng YM. (2023)Research on Online Course Design of Information Technology Based on Subject Knowledge Graph [J]. Chinese Journal of IC T in Education, 2023,29 (08): 121-128.
2. Lang YK, Su C, Wang GZ. (2021)The Construction and Reasoning of C++ Course Knowledge Graph Based on Neo4j [J]. Intelligent Computer and Applications, 2021,11 (07): 144-150+155.
3. Zhang HN, Song Y. (2023)Construction of Course Knowledge Graph for Smart Education [J]. Computer Education, 2023 (09): 120-125. DOI: 10.16512/j.cnki.jsjy.2023.09.09.028.
4. Dilinur K í mu. (2022)A Study on the Importance of Constructing a Course Knowledge Graph - Taking the "Data Structure" Course as an Example [J]. Electronic Components and Information Technology, 2022,6 (07): 239-242. DOI: 10.19772/j.cnki.2096-4455.2022.7060.
5. Cao HW, Xu JL, Dou FK. (2020)Building Biomedical Knowledge Graph with Neo4j Database [J]. Computer Era, 2020 (06): 35-38. DOI: 10.16644/j.cnki.cn33-.
6. Sun F. (2020)Application of Knowledge Graph in Financial Institution Network Security [J]. Financial Technology Time, 2020,28 (06): 82-85.
7. Xiu LQ. (2023)Online Information Literacy Education Knowledge Map Construction and Recommendation System Implementation [D]. Beijing University of Civil Engineering and Architecture, 2023. DOI: 10.26943/d.cnki.gbjzc.2022.000568.
8. Liu XL, Wang W. (2024)The course teaching reform based on knowledge map [J]. China metallurgical education, 2024,(01):7-11. DOI: 10.16312 / j.carol carroll nki/g4.2024.01.004 cn11-3775.
9. Wei Y, Cui GX.(2024)The Python program design course teaching idea based on knowledge map [J]. Journal of computer education, 2024, (02):51-54. DOI: 10.16512 / j.carol carroll nki jsjjy. 2024.02.017.
10. Yan H. (2024)Research on the construction and application of Curriculum Knowledge Graph in Blended teaching [J]. Computer knowledge and technology, 2024,20 (02):175-177. The DOI: 10.14004 / j.carol carroll nki CKT. 2024.0007.