Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 16, No. 1, 2025

The Construction and Application Practice of Course Knowledge Graph in Educational Digitalization

Jian Tian

School of Physical Education Science, Fujian Normal University, Fuzhou Fujian, 35000, China

Abstract: As the infrastructure for semantic integration modeling of course knowledge points, learning tasks and evaluation indicators, the Curriculum Knowledge Graph (CKG) has become a key fulcrum for educational digitalization. Under the theoretical framework of networked learning and knowledge organization, this paper defines the conceptual boundary and mechanism of CKG, and summarizes the engineering processes such as ontology modeling, information extraction, entity alignment, knowledge warehousing and reasoning. Further integrating application evidence such as instructional design optimization, personalized path generation, and enhanced generation of learning situation diagnosis and retrieval (RAG), a quality assessment system and versified governance approach centered on structural consistency, coverage completeness, usability benefits, and trusted traceability are proposed. And discuss the synergy of large language models in term discovery, relational assumptions, and weakly supervised annotations, as well as the auditability gains and consistency risk they bring.

Keywords: Curriculum Knowledge; Graph Ontology Construction; Personalized Learning; Quality Assessment.

1. Introduction

Educational digitalization is moving from "resource digitalization" to the deep-water zone of "data and knowledge intelligence". As a result, the structuring, semanticization and computability of courses have become the key infrastructure supporting the modernization of teaching [1]. Existing research has long been biased towards the single-modal nature of text and has rarely integrated multi-modal evidence such as classroom records, experimental processes, source codes, and question banks. This can easily lead to problems such as knowledge fragmentation, untraceable prior dependencies, and insufficient evidence for teaching decisions. With the Course Knowledge Graph (CKG) as the common foundation, teaching objectives, course contents, learning activities and evaluation methods can be uniformly mapped into a retrievable, inferable and visualizable semantic network. Thus, at the practical level, the collaborative effectiveness of personalized learning, learning analysis and educational resource governance can be simultaneously enhanced. Compared with the traditional "course teaching" centered paradigm, knowledge graphs take entities, relationships, and levels as the core modeling units, enabling the mapping between goals and activities to have a formal and auditable expression, and thereby supporting cross-course transfer and cross-version iteration. Under the joint influence of policies and industries, the continuous revision of curriculum standards, the large-scale growth of resource platforms, and the online accumulation of learning data have further given rise to the practical demand for an explainable, evolvable, and governable knowledge infrastructure to reduce the marginal costs of heterogeneous resource integration and continuous updates. Based on recent reviews and practices, it can be observed that the engineering paradigm centered on construction, evaluation, and application is gradually taking shape, and multimodal and deep learning technologies are becoming key driving forces [2] especially showing significant progress from pilot projects to maturity in terms of evidence alignment, cross-modal consistency, and teacher-end usability.

2. Concept Definition and Theoretical Basis

The curriculum knowledge graph takes curriculum standards, knowledge points, ability and quality, learning tasks and evaluation indicators and their interrelationships as the basic units[3]. On this basis, using the "core course knowledge graph" to organize and construct the content system of teaching materials in reverse is becoming a new practical path. Its fundamental goal is to make the course knowledge and learning path computable and interpretable. Compared with the general encyclopedia graph with a wider coverage and also different from the concept map and mind map with weaker semantic constraints, the course knowledge graph places more emphasis on achieving the consistency of knowledge and strict semantic association under the constraints of ontology.

In terms of theoretical support, constructivism and associationism view learning as a continuous reconstruction process within "nodes, connections, and networks", which naturally aligns with the structural expression of knowledge graphs: Learners achieve knowledge transfer and meaning construction by forming and strengthening key connections in the network, while teachers optimize the teaching pace and strategies around key nodes (such as core concepts and bottleneck knowledge) and key edges (such as prerequisite dependencies and error-prone associations), to realize the isomorphic mapping of goals, content, activities and evaluations[4]. Furthermore, the course knowledge graph is coupled with learning analytics and educational measurement theories: the ontology (or model) layer provides a semantic coordinate system for the measurement dimensions, the data layer carries evidence such as assignments, tests, and interactions, and the reasoning layer updates to specific knowledge units based on the evidence, thereby forming a traceable and auditable representation of mastery.

3. Construction Methods and Processes

The construction of a course knowledge graph can be

summarized as a five-link process that is interrelated: ontology modeling, information extraction, entity alignment and disambiguation, database entry and reasoning, and service and maintenance [5]. This process serves both as a methodological path and an engineering governance framework, emphasizing the goals of standardization, interpretability, and evolwability, and gradually implementing it from the conceptual layer to the data layer and then to the application layer.

(One) Ontology construction. Taking the curriculum standards and teaching syllabus as the anchor points, complete the systematic design of term collection, concept stratification, and relationships and constraints, and achieve semantic alignment of "teaching objectives - teaching content - ability requirements - evaluation methods" through consistency verification. The fundamental principles of ontology engineering are lightness, reusability and scalability.

(Two) Information extraction. For multi-source data such as teaching materials, lesson plans, question banks, classroom and experiment logs, code repositories and experiment reports, a comprehensive approach is adopted, including entity recognition, relationship and event extraction, as well as template matching and other technical routes.

(Three) Alignment and disambiguation to address conceptual conflicts of the same name and synonym, cross-version and cross-resource, a stable namespace and mapping relationship should be constructed, and at the same time, a conflict resolution and change auditing mechanism should be established. For different definitions of the same term, differentiated definitions of equivalent relationships, and inconsistent superior-subordinate relationships, it is necessary to form traceable change records and rollback strategies. When aligning versions, semantic differences and their scope of influence should be explicitly recorded to prevent hidden drift from causing deviations in subsequent teaching analysis and application.

(Four) Warehousing and reasoning. According to application requirements, a trade-off should be made between the semantic Web and property graph paradigms: using RDF and OWL is convenient for implementing rule constraints and consistent reasoning, while using property graphs (such as Neo4j) is convenient for efficient relationship traversal and business system integration. For classroom behaviors and teaching events, modeling can be carried out through event and time series expansion to support computational tasks such as learning path generation, trajectory analysis, and risk early warning.

(Five) Service and Maintenance. Provide integrated services for teaching around retrieval, question answering, recommendation, learning profiling and visualization, and incorporate data calibration, crowdsourcing review, version release and log auditing into the knowledge engineering process of continuous integration and continuous delivery. Provide teachers with an easy-to-use graph editing and proofreading workbench, and offer students an explainable learning panel with evidence chains and path explanations.

4. Evaluation of Teaching Application and Practical Effect

At the level of instructional design, the course knowledge graph can consistently map teaching objectives, teaching content, learning activities and evaluation elements, enabling teachers to identify the parallel and extended relationships between key contents and knowledge points, optimize the teaching sequence, and reduce the risks of objective deviation and resource misallocation. For cross-class or cross-version collaborative lesson preparation, the knowledge graph, based on a shared semantic foundation, provides path adjustment suggestions of "minimizing differences" to ensure consistency and transferability in core knowledge and ability requirements across different teaching scenarios.

At the level of personalized learning and resource recommendation, the knowledge graph projects the dependency structure of knowledge points and the mastery degree of learners into a unified semantic space, comprehensively considering the difficulty and relevance to generate personalized paths for each student, and pushing corresponding practice and remedial resources [6]; Its interpretability stems from the explicit presentation of the dependency link, which not only facilitates teachers' review but also helps enhance students' learning motivation. At the level of learning situation diagnosis and formative assessment, the knowledge graph maps homework, tests and interaction behaviors to the dimensions of knowledge and ability, supporting a closed-loop process of "weak concept location, error pattern attribution, strategic intervention and effect retesting". For teaching administrators, the multi-dimensional portrait of "class - teacher - course - resource" based on the graph can provide auditable evidence for quality monitoring and resource allocation.

At the level of resource retrieval and question answering, knowledge graphs enhance the reliability of retrieval and generation through semantic constraints and evidence alignment, reducing semantic ambiguity and answer illusions caused by free text retrieval. At the level of classroom intervention, the combination of knowledge graphs with online logs and sensor data can identify high-risk concepts and learning stagnation points, and push practical intervention suggestions to teachers. In terms of efficacy evaluation, it is recommended that the pre - and post-test and control groups be taken as the basic design units, supplemented by teacher interviews and learning log analysis. This will comprehensively examine indicators such as learning effectiveness, learning efficiency, resource coverage and hit, teacher workload, interpretability and satisfaction. At the same time, internal validity threats such as teacher expectation effect, historical effect and selection bias should be actively controlled in the experimental design and data analysis. Existing reviews indicate that advancing the evaluation design and engineering process in parallel can help address both technical availability and teaching effectiveness issues simultaneously during the implementation stage, significantly reducing promotion resistance and enhancing the robustness of large-scale implementation.

5. Quality Assessment and Course Knowledge Graph Governance

The quality of a course knowledge graph is not only reflected in its correctness, completeness and consistency, but also should highlight the overall synergy of usability, credibility and evolwability. In terms of structural quality, ontology consistency checking, constraint satisfaction checking and error pattern scanning should be incorporated into the regular process to systematically address typical defects such as duplicate entries, semantic conflicts, link interruptions and isolated subgraphs. In terms of coverage

quality, the curriculum standards and examination syllabi should be taken as the reference system. At the same time, the coverage level of nodes and relationships, the continuity of the prerequisite chain, and the coverage differences among different versions should be measured to avoid "false saturation" caused by the absence of key prerequisite items when the surface coverage is sufficient. At the usability level, the evaluation should take task gain as the core indicator, comprehensively observe the improvement in recommendation, retrieval and prediction compared to the baseline, and incorporate interpretability and teacher usability into the performance judgment. At the level of credibility, emphasis should be placed on source labeling and evidence chain recording, along with conflict resolution and change traceability mechanisms, to ensure that every key relationship has an auditable source and a clear change history.

To transform quality objectives into executable engineering paths, the evaluation process should adopt a combined paradigm of "small-scale gold standards, sampling manual review and online control experiments" : the gold standard is used to anchor the accuracy of key facts, sampling review is used to control pipeline deviation and annotation drift, and online experiments are used to test the gain and robustness in real teaching scenarios. The corresponding governance mechanism should embed version management, semantic change tracking, role and permission management, proofreading process design, as well as privacy and copyright compliance into the build pipeline [7]. Version upgrades should be accompanied by impact analysis and automated regression testing. During the review and proofreading process, it is necessary to distinguish between semantic modifications and data patches. For resources involving privacy or protected by copyright, verifiable compliance templates should be provided in terms of usage scenarios and output forms. A review of relevant technologies has shown that key technologies, application values, and challenge lists are still evolving in the context of deep learning and large models. Therefore, quality control and governance design should not be postponed but should be carried out in parallel with the construction process from the very beginning of the project, forming a closed loop from construction to evaluation and then to application and governance, thereby supporting the sustainable evolution and large-scale implementation of knowledge graphs.

6. Major language Models and Their Future Development directions

Large language models bring significant benefits and controllable risks to course knowledge graphs, and an overall design at the system architecture level is required. In the construction stage of human-machine collaboration, the model can undertake tasks such as term discovery, relationship hypothesis generation, and weakly supervised annotation suggestions, thereby providing effective automation gains beyond high-cost manual operations. If rule constraints and generative methods are connected in parallel as a dual-track process, the former provides high-confidence verifiable priors, while the latter is responsible for covering the long tail and completing candidates. The two form a closed loop through consistency check and error reinjection, which can significantly reduce the marginal cost of cold start and large-scale updates. Meanwhile, the semantic drift and consistency risks caused by generative uncertainty cannot be ignored and must be suppressed with a complete set of "guardian mechanisms" [8]. Specifically, the prompt engineering and output templates provide structural constraints for the generated results; Ontology consistency verification ensures that the generated content is compatible with existing semantic constraints. Evidence retrieval alignment is used to compare the correspondence between generated fragments and knowledge sources; Version governance linkage is used to trigger alerts and regression tests in a timely manner during the evolution of the graph. Through the above multi-layer constraints, the generation advantages of the model can be mutually reinforced with the structural prior of the knowledge graph within the controllable boundary.

The development path for the future can be advanced from four dimensions: First, build a process-oriented curriculum map covering learning behaviors, teaching events and evaluation results, so that classroom dynamics and strategic interventions can be computationally expressed in the time dimension. Secondly, establish a bidirectional alignment interface between the graph and the model, use the ontology fragments as prompt templates, and simultaneously screen and reorder the generated candidates based on graph constraints to form a stable structure-generation closed loop. Thirdly, accumulate cross-disciplinary and reusable upper-level ontologies and competency models, and provide cross-domain alignment and migration tools to reduce the resource consumption of repetitive modeling. Fourth, jointly build public benchmarks and consensus indicators with industry, academia and research institutions, provide comparable protocols and open-source samples at the three levels of construction, evaluation and application, and promote the construction of a standardized ecosystem for large-scale and sustainable implementation. The common goal of the above agenda is to enable knowledge engineering and model engineering to form a long-term co-evolution under the constraints of governance and evaluation. Knowledge graphs provide controllable structures and evidence, while language models offer efficient generation and automation. Both continuously iterate within a unified quality and compliance framework, ultimately achieving reliable support for the teaching and learning process.

7. Conclusion

The core of the course knowledge graph lies in integrating scattered course objectives, knowledge points, teaching activities and evaluation indicators into a governable knowledge network based on explainable structural priors and computable reasoning mechanisms, thereby promoting the teaching paradigm to gradually shift from being experience-dominated to being driven by both evidence and mechanisms. This article, following a logical thread of six parts, systematically sorts out the relevant concepts and theoretical context, summarizes the engineering construction process, generalizes typical teaching application scenarios, summarizes the key points of quality control and governance, and discusses the latest progress and risk boundaries of the coordinated development with large language models. Overall, the value of the course knowledge graph is expanding from merely optimizing resource organization to enhancing support for teaching decisions, visualizing the learning process, and facilitating the trusted circulation of educational data elements. For practical implementation, it is advisable to take the lightweight and reusable course itself as

the anchor point, and adopt a weakly supervised extraction strategy and versioned governance process to establish a continuous evaluation closed loop composed of small-scale gold standards, sampling reviews and online experiments. At the system architecture level, a composite paradigm combining graph constraints and retrieval-enhanced generation is adopted to improve auditability with structured evidence, while leveraging the advantages of generative models in content generation and efficiency. For medium and long-term development, the roadmap should incorporate event and process-oriented course maps, the bidirectional alignment mechanism of maps and models, and the co-construction of cross-disciplinary upper-level ontologies. Moreover, through open data and consensus indicators, the comparability and large-scale replication capabilities across institutions should be enhanced. Ultimately, an intelligent infrastructure will be formed that can not only reduce the burden on teachers and increase their efficiency, but also enable learners to teach students in accordance with their aptitudes, and provide measurable and auditable support for educational governance.

References

[1] Hu Haibin, Ding Guozhu, Wu Pengfei. Construction and Application of Course Knowledge Ontology Based on Learning Meta-Platform: Taking the "New Development of Educational Technology" Course as an Example [J] Research on Educational Technology,2017,38(10):74-81.

[2] Gao MAO, Zhang Liping. Research on the Connotation, Technology and Application of Educational Knowledge Graph Integrating Multimodal Resources [J]. Application Research of Computers,2022,39(08):2257-2267.

[3] Tan Fangzheng, Zhang Long, Shan Lei, et al. The Chinese Model and Practical Path for the Construction of "101 Plan" Textbooks in the Basic Discipline Series [J] Chinese Editor,2025, (09):4-9.

[4] Yao Lin, Ma Pengfei, Shang Danmei, Jiang Yong. Research on the Construction of Subject Knowledge Graph Based on Python Programming [J]. Chinese Journal of Management Informatization, 2020,28(12):127-130.

[5] Shi Jiangyong, Tang Jintao, Wang Yongjun, Zhang Long. Construction of Teaching Resources for Emerging Field Courses Based on Knowledge Graph [J]. Higher Engineering Education Research,2022(03):15-20.

[6] Zhao Yubo, Zhang Liping, Yan Sheng, Hou Min, Gao MAO. Review on the Construction and Application of Subject Knowledge Graphs in Personalized Learning [J]. Computer Engineering and Applications,2023, 59(10):1-21.

[7] Sun Lijun, Meng Fanjun, Xu Xingjian. A Review on the Construction Technology of Course Knowledge Graph [J/OL]. Computer Engineering,1-25[2025-11-08].

[8] Chen Gang, Sun Yuantao. Knowledge "Supply" : How Human-Machine Collaborative Knowledge Production Can Drive "Dual Intelligence" [J]. Research on Educational Technology, 2020,46(11):25-32.