

Research on Multi Source Data Fusion Evaluation System Based on Modular Knowledge Graph: Taking the Course of Database Principles and Applications in Colleges and Universities as an Example

Haoyue Lu, Pengfei Su

Sias University, Zhengzhou, Henan 450000, China

Abstract: In view of the pain points of single evaluation, lack of process supervision, and weak practical evaluation of the course "Database Principles and Applications" in Colleges and universities, this study constructed a smart evaluation system based on modular knowledge map and integrating multi-source data (online behavior, experimental reports, project results, and enterprise evaluation), so as to achieve accurate portrait in the whole learning cycle. After a semester of controlled experiment (242 students), the experimental group was significantly better than the control group in the level of knowledge structure, engineering practice ability and autonomous learning engagement ($p < 0.05$), which effectively solved the dilemma of "emphasizing results over process, emphasizing theory over practice". The system provides a replicable practice paradigm for the digital and accurate evaluation of computer courses in local private colleges and universities.

Keywords: Modular Knowledge Map; Principle and Application of Database; Multi Source Data Fusion; Process Evaluation.

1. Introduction

1.1. Research Background

Under the background of the comprehensive promotion of the education digitalization strategy, the opinions on accelerating the promotion of education digitalization issued by the Ministry of education and other nine departments clearly put forward the idea of "promoting the construction of subject knowledge map", emphasizing the innovation of teaching evaluation enabled by emerging technologies. Database principle and application, as the core basic course of computer specialty, has the characteristics of high knowledge density, strong logical correlation and outstanding engineering practice attribute, which is the key bridge connecting theory and application. However, at present, the course generally follows the traditional evaluation mode of "usual score+final written examination", and there are three core pain points: first, the emphasis on results and the neglect of process, and the lack of process learning supervision caused by the dominance of final score; Second, it emphasizes memory and ignores structure, and the fragmented evaluation of knowledge points is difficult to reflect the overall construction of knowledge system; Third, it pays more attention to theory than practice, and the evaluation of project design, code implementation and other engineering capabilities is insufficient, and the evaluation criteria are out of line with the job requirements of enterprises.

Knowledge mapping technology can integrate scattered knowledge points into a dynamically related knowledge network, which has the advantages of structure and visualization; The modular design supports the flexible splitting and reorganization of knowledge units, and can adapt to the hierarchical evaluation needs of different teaching stages. The combination of the two provides an ideal underlying framework for the reform of curriculum evaluation. Based on this, this study integrates multi-dimensional data such as knowledge mastery, learning

behavior, practical tasks, and enterprise evaluation, and constructs an accurate evaluation system covering the whole process of "before class - in class - after class" and running through the full dimension of "knowledge ability literacy", aiming to fundamentally reform the traditional curriculum evaluation mode and provide reusable practice paradigm for similar curriculum reform.

1.2. Research Significance

At the theoretical level, this study breaks through the limitations of the traditional application of knowledge mapping in Resource Recommendation and learning path planning, extends it to the field of curriculum evaluation, and constructs the evaluation theoretical framework of "modular knowledge mapping+multi-source data fusion", which enriches the theoretical system of process and accurate evaluation from the perspective of smart education. At the practical level, aiming at the persistent disease of "emphasizing results over process, emphasizing theory over practice" in the database course evaluation, this study designed and verified a practical wisdom evaluation scheme to achieve the dual goals of "promoting learning and teaching through evaluation". In addition, the lightweight and standardized evaluation resource package has strong portability, which can effectively reduce the threshold of technology application and provide a feasible path for the inclusive improvement of digital teaching ability of local private colleges and universities.

1.3. Research Content and Methodology

This study focuses on four aspects of work centered around the "evaluation system for multi-source data fusion enabled by modular knowledge graph": First, construct a modular knowledge graph for "Database Principles and Applications", dividing modules according to "basic - advanced - comprehensive - expanded" to form reusable knowledge anchors; second, design an evaluation index system for multi-source data fusion, integrating four types of data: online

behavior, classroom performance, project practice, and enterprise evaluation, and using the Analytic Hierarchy Process (AHP) to determine weights; third, establish a closed-loop operation mechanism of "graph anchoring - data collection - fusion calculation - feedback intervention"; fourth, conduct a semester of teaching practice, comparing multi-dimensional data between the experimental class and the control class to verify the effectiveness of the system.

Research methods include: literature research method (combining knowledge map and evaluation theory), case study method (using typical PBL projects as the carrier), controlled experiment method (comparing experimental class and control class, using t-test), and data analysis method (cleaning, standardizing, and conducting correlation analysis on multi-source data).

2. Related Research Foundation

2.1. Current Application Status of Knowledge Graph in the Field of Education

The knowledge graph organizes knowledge in a structured manner of "node-relationship-attribute", which can clearly present the logical connections between knowledge points. In the context of the continuous advancement of educational informatization, constructing a precise and efficient curriculum knowledge graph has become one of the key tasks in promoting personalized education development. Currently, the application of knowledge graphs in the field of education mainly focuses on scenarios such as personalized learning path recommendation, intelligent question-answering systems, and curriculum resource organization. For example, Ma Ruixin et al. [1] constructed a teaching resource system covering course knowledge points and ability graphs based on the digital textbook "Database Principles and Applications: Based on OpenGauss". They provided 13 modular experiments on Huawei cloud hosts, realizing a smart learning mode of "learning while practicing". Fan Qili et al. [2] explored the application of "AI + knowledge graph" in the "openGauss Database Technology" course in vocational colleges. They constructed a curriculum knowledge graph combined with artificial intelligence technology for personalized and precise teaching, aiming to empower the construction of "golden courses" and teaching reform. In addition, the School of Food Science and Engineering at Jiangnan University constructed an AI teaching system based on the knowledge graphs of 20 core courses, achieving 100% professional coverage. After the system pushed personalized learning plans, the average knowledge retention rate of students jumped from 58% to 82%. However, existing research mostly focuses on the macro construction of the overall knowledge network, and the design of modular, separable, and reusable graphs for database courses is still inadequate. There is still much room for exploration in practical needs such as hierarchical teaching, phased assessment, and flexible teaching organization.

2.2. Research Progress on University Curriculum Evaluation System

In recent years, the evaluation of courses in domestic universities has gradually shifted from summative evaluation to process-oriented and diversified evaluation. Scholars have actively explored the reform of course assessment models in the context of "Internet plus" and proposed ideas for constructing diversified process-oriented assessment models.

Some studies, set against the backdrop of applied undergraduate talent cultivation, have explored the formation of a whole-process and diversified course assessment model through the optimization of course evaluation components and learning evaluation dimensions, effectively leveraging the role of "assessment to promote learning and teaching". In terms of technology-enabled evaluation, the "Xidian Smart Assessment" case launched by Xidian University [3] has been selected as a typical application scenario for "Artificial Intelligence + Higher Education" by the Ministry of Education. This case focuses on the three goals of "personalized cultivation, intelligent adaptation, and comprehensive development", and has constructed a five-in-one closed-loop feedback system for student evaluation chain, including "electronic archives - digital portrait - comprehensive assessment - competency certificate - growth assistant", promoting the transformation of student evaluation from a management tool to an educational ecosystem. North China University of Water Conservancy and Electric Power has taken multi-source teaching data governance as a breakthrough point, constructing a data collection network covering the entire teaching process, effectively breaking the data silos between the teaching management system and smart classrooms, integrating teaching basic services, resource management, teaching implementation, and assessment evaluation functions, and achieving full-process data connectivity and multi-scenario collaboration. Henan University of Technology [4] has constructed a new paradigm of university teaching governance through the construction of an AI evaluation system, featuring "data-driven, human-machine collaboration, and closed-loop improvement", forming a normalized classroom monitoring mechanism where full-course AI evaluation replaces periodic inspections. It is noteworthy that some scholars have introduced knowledge graphs into the field of course evaluation, proposing a knowledge graph-based big data analysis model for course quality, aiming to address prominent issues such as semantic bias, false evaluations, and information distortion in online education course evaluation [5]. However, most universities' evaluation reforms still remain at the level of combined adjustments such as "attendance + homework + experiment + examination", and the deep integration and intelligent analysis of multi-source data have not been fully realized. The granularity of evaluation results is relatively coarse, and the accuracy needs to be improved, making it difficult to effectively support the practical needs of individualized teaching and precise instruction.

2.3. The Practical Dilemmas of Database Course Teaching and Evaluation

In the specific practice of database courses, the academic community has actively explored teaching reform and evaluation optimization. Ma Ruixin and others have constructed digital textbooks using the domestic database OpenGauss as a practical platform, organizing database-related knowledge points into 13 modular experiments through the setting of a capability map, reflecting the modular design concept of the database course knowledge system. Lu Yanjing, Fan Qili, and others have explored a personalized and precise teaching model driven by "AI + knowledge map" for the "openGauss Database Technology" course, aiming to promote the improvement of teaching quality in database courses. In the field of private universities, researchers have implemented "three excellences and three newness" teaching

innovation and practice for the database principles and applications course in private applied undergraduate colleges and universities in the context of information and communication technology (ICT) [6]. Starting from three dimensions: optimizing course content to reconstruct new forms, optimizing teaching methods to deepen new educational reforms, and optimizing evaluation modes to enhance new quality, the aim is to improve the teaching effectiveness and talent cultivation quality of database courses. However, the aforementioned research is still relatively scattered and has not yet formed a systematic evaluation reform plan. Overall, in the group of private universities, database courses generally face practical problems such as lagging textbook updates, insufficient experimental resources, uneven information technology capabilities among teachers, and rigid evaluation methods. Students often exhibit phenomena such as fragmented knowledge, weak practical operation abilities, and insufficient learning interest during the learning process, urgently requiring the systematic reconstruction of teaching processes and evaluation systems with the help of emerging technological means [7].

3. Construction and Design of Modular Knowledge Graph

3.1. Design Concept of Modular Knowledge Graph

Traditional knowledge graphs often adopt a "global single network" construction model, incorporating all knowledge points into a highly coupled network structure at once. Although this approach helps present a comprehensive view

of knowledge, it reveals issues such as insufficient flexibility, poor reusability, and difficulty adapting to hierarchical teaching and periodic evaluation needs in practical teaching applications. To address this, this study breaks through the limitations of the holistic graph and proposes a design concept of "general-specific combination and modular decomposition".

The core idea is as follows: Firstly, based on the course syllabus and typical textbooks, all the knowledge points of "Database Principles and Applications" are divided into several relatively independent knowledge modules according to their inherent logical relationships. Secondly, within each module, a detailed sub-graph is constructed to clarify the dependencies, progression, and associations between nodes. Finally, the overall connection is achieved through the interface relationships between modules (such as pre-knowledge references and subsequent extension links), forming a flexible knowledge network architecture with "high cohesion within modules and low coupling between modules". This design supports flexible invocation by chapters, knowledge points, or practical units, and can adapt to the evaluation needs of classroom teaching, experimental training, project practice, final review, and other scenarios. It provides dynamically configurable knowledge anchors for subsequent multi-source data fusion evaluation.

3.2. Modular Division of Courses

Based on a systematic review of the knowledge system of the "Database Principles and Applications" course, and taking into account both students' cognitive patterns and engineering practice logic, this study divides the course into six independent and reusable knowledge modules, as detailed in Table 1 below:

Table 1. Knowledge Module Table

Module Number	Module Name	Core Content	Example of Main Knowledge Points
Module 1	Basic theory of database	Basic concepts of database, data model, and database system structure	Data abstraction, three-level schema with two-level mapping, database language, data independence
Module 2	Relational database and relational algebra	Relational model, relational algebra operations	relation definition, selection, projection, join, division operation, tuple relational calculus
Module 3	SQL Language and Query Optimization	SQL DDL/DML/DCL, views, indexes, query optimization	Table creation, CRUD (Create, Read, Update, and Delete), nested queries, and aggregate functions
Module4	Database design (ER model, normalization)	Conceptual design, logical design, normalization theory	ER diagram, primary key and foreign key, 1NF to BCNF, schema decomposition, design tool
Module 5	Transactions and concurrency control	ACID properties, isolation levels, and locking mechanisms of transactions	Transaction commit rollback, dirty read and non-repeatable read, two-phase locking protocol, deadlock handling
Module 6	Database security and operation and maintenance	Permission management, backup and recovery, performance monitoring	User authorization, role management, log backup, data recovery, and slow query analysis

The six knowledge modules are: Basic Database Theory Module, Relational Database and Relational Algebra Module, SQL Language and Query Optimization Module, Database Design (ER Model, Normal Forms) Module, Transaction and Concurrency Control Module, and Database Security and Operation and Maintenance Module. This division covers the core content of the course and maintains the relative independence and inherent logical connection between the modules.

3.3. Implementation of Knowledge Graph Technology

To achieve the aforementioned modular design, this study utilizes the Python programming language in conjunction with the Neo4j graph database to construct and store the knowledge graph.

(1) Node design: Each knowledge node corresponds to a fine-grained knowledge point. A total of more than 150 nodes have been extracted and defined, covering five categories: core concepts (such as "relation" and "normal form"), principles (such as "two-phase locking protocol"), syntax (such as "SELECT statement"), methods (such as "schema decomposition"), and skill points (such as "database backup operation"). Each node is assigned a unique ID for easy binding in subsequent evaluation tasks.

(2) Relationship edge design: Nodes are connected through directed edges, defining a total of 6 types of relationships, with over 200 edges in total. Figure 1 shows the relationship edge design of some nodes. These include inclusion relationships (where a larger concept includes sub-concepts, such as "database design" → ER model); dependency relationships (where learning B must precede learning A, such

as "normal form" dependent on "functional dependency"); progressive relationships (where there is a vertical increase in knowledge difficulty or depth, such as "basic SQL" → advanced query); application relationships (where theoretical knowledge points are used for practical skills, such as "select projection" → SQL implementation); association relationships (where there is a horizontal connection between knowledge points without a strict order, such as "transaction" and "recovery"); and constraint relationships (where one knowledge point restricts the implementation method of another, such as "normal form constraint" affecting "table splitting").

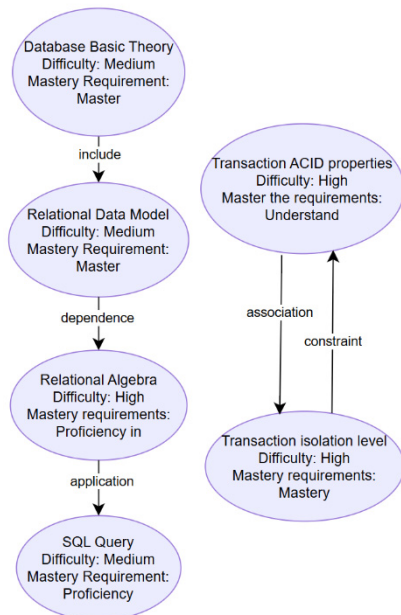


Fig 1. Example diagram of some nodes

(3) Attribute label design: Attach multi-dimensional attribute labels to each node to support precise evaluation and learning diagnosis. Difficulty level: Beginner, Intermediate, and Advanced, corresponding to the three cognitive levels of memorization, comprehension, and application respectively. Mastery requirements: Divided into "Understand", "Comprehend", "Master", and "Proficiently Master" according to the teaching syllabus. Associated practical tasks:

Point to specific experimental or project task IDs. Assessment frequency: Record the number of times the node appears in various quizzes and exams. Weak spot marking: Dynamically identify knowledge points that have been frequently made mistakes or are difficult to understand by previous students, as shown in Figure 2.

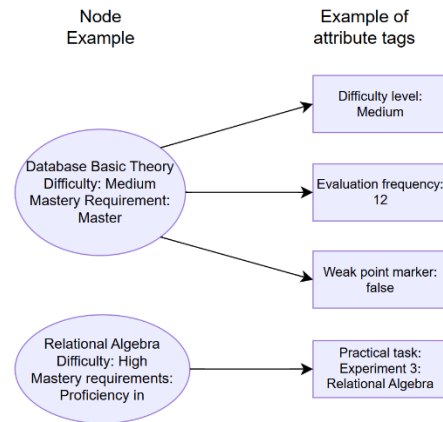


Fig 2. Example of attribute labels

In terms of technical implementation, firstly, Python is utilized to develop crawlers and text parsing tools, extracting knowledge points and their relationships from course materials, lesson plans, and past examination questions. Then, Pandas is employed for data cleaning and structuring. Finally, Neo4j's Cypher query language is used to batch import nodes and relationships, constructing a visual graph. This graph supports various functions such as visual knowledge network display, unidirectional or bidirectional tracing of knowledge points (e.g., predecessor-successor queries), association queries (e.g., "find all nodes related to SQL optimization"), and automatic identification of weak spots (based on high-frequency error nodes annotated from historical assessment data). It provides precise data anchors and knowledge basis for the evaluation system.

4. Construction of Evaluation System for Multi-Source Data Fusion

4.1. Multi-Source Data Dimension Design

Table 2. Multi-source data dimension design

Data Category	Sub-dimension	Specific collection content	Frequency
Knowledge mastery data	Knowledge point testing, concept differentiation, graph node traversal depth, and knowledge association accuracy	Scores of objective questions in chapter quizzes and mid-term tests Accuracy of judging similar concepts (views, tables, etc.) Node clicks, dwell time, and secondary query counts Correct rate of filling in the blanks for antecedent/successor/relational relationships	Each chapter weekly real-time Each module
Process behavior data	Classroom interaction frequency Experimental submission deadline Participation in group discussions Platform learning duration Quality of homework completion	Number of participation in questioning, voting, bullet comments, etc On time/early/delayed days Number of forum posts, replies, and adoption Total duration of videos, documents, and exercises Objective question score+teacher's gauge rating	Each lesson Every experiment weekly daily weekly
Practical task data	Completion of PBL project SQL code standardization Query optimization capability Reasonable database design	The completeness of deliverables at each stage Indentation, Annotation, Naming Rules Rating Comparison of execution efficiency before and after optimization ER diagram, paradigm, index rationality	project lifecycle Project submission Experimental class hours project review
Evaluation data of industry education integration	Enterprise Mentor Rating Workplace adaptability	Innovation, document standardization, and application value Understanding enterprise processes and using collaboration tools	Mid term+End of term final

This study integrates four major types of core data to construct a comprehensive and multi-dimensional evaluation model. Specifically, as shown in Table 2, the first type is knowledge mastery data, including knowledge point testing, concept differentiation, graph node access depth, and accuracy of knowledge association. The second type is process behavior data, including classroom interaction frequency, experimental submission timeliness, group discussion participation, platform learning duration, and homework completion quality. The third type is practical task data, including PBL project completion degree, SQL code normalization, query optimization ability, and database design rationality. The fourth type is industry-education integration evaluation data, including scores given by enterprise mentors on project plans, engineering specifications, application value, and workplace adaptability.

To ensure the integrity and authenticity of the data, all automatically collected data undergoes outlier detection (such as removing invalid clicks below 30 seconds), and teacher and enterprise mentor ratings are conducted in a double-blind back-to-back manner to reduce interference.

4.2. Evaluation Indicators and Weight Allocation

In the evaluation index system shown in Table 3, the first level indicators include knowledge mastery level, learning process performance, engineering practice ability, and professional comprehensive literacy, with weights of 30%, 25%, 30%, and 15%, respectively. The level of knowledge mastery is divided into secondary indicators such as concept understanding, relationship analysis, and knowledge point coverage. The data is sourced from online testing and knowledge graph access data. The learning process performance includes secondary indicators such as classroom participation, submission timeliness, self-directed learning, collaborative discussions, etc. The data is sourced from the behavior records of the learning platform. The engineering practice capability includes secondary indicators such as project completion, code standardization, query optimization, and system design, with data sourced from experimental reports and PBL project results. Vocational comprehensive literacy includes secondary indicators such as team collaboration, engineering standards, industry adaptation, and innovation awareness, with data sourced from group peer evaluations and enterprise mentor ratings.

Table 3. Evaluation Index System

First-level indicator	Weight	Secondary indicator	Weight (occupying one level)	Data source
Knowledge mastery level	30%	Conceptual Understanding	40%	Online testing (objective questions)
		relationship analysis	35%	Knowledge graph related fill in the blank and concept comparison questions
		Knowledge point coverage rate	25%	Graph access log (number of access nodes/total number of nodes)
Performance during the learning process	25%	class participation	30%	Classroom interaction frequency+in class quizzes
		Submission deadline	20%	LMS homework submission timestamp
		self-directed learning	30%	Platform learning duration+video completion rate
		Collaborative discussion	20%	Forum activity+group mutual evaluation contribution score
Engineering practical ability	30%	Project completed	30%	Integrity of deliverables at each stage of PBL
		Code Standards	25%	Enterprise mentors score based on standardized checklists
		query optimization	25%	Compare the efficiency of online evaluation systems before and after optimization
		System Design	20%	ER graph quality+paradigm level+index
Comprehensive professional competence	15%	team collaboration	30%	Group mutual evaluation+enterprise mentor observation
		Engineering specifications	25%	Document structure and version control usage
		industry adaptation	25%	Enterprise mentor evaluates the application value of the plan
		Innovation awareness	20%	Innovation in technology selection and exception handling in the project

The above weights can be fine tuned according to the semester's curriculum priorities (such as a semester's emphasis on SQL experiments), but the adjustment range should not exceed $\pm 5\%$, and should be discussed and confirmed by the teaching team to ensure the stability and comparability of the evaluation.

4.3. Evaluation and Implementation of Closed-loop Process

The operation of the evaluation system follows the five step closed-loop logic of "anchoring acquisition calculation feedback iteration", forming a spiral process from goal setting to continuous improvement. The following figure shows the process (described in words here, flow chart can be provided in the actual paper).

Step 1: Map anchoring. Before each teaching activity

(classroom test, experiment, PBL milestone), teachers check the knowledge map node assessed by this activity in the teaching platform. For example, a group task on "Index Optimization" will bind nodes such as "B+Tree Index" and "Execution Plan Analysis" in Module 3. The system automatically records the binding relationship to provide a basis for subsequent data attribution.

Step 2: Multi dimension collection. The platform automatically collects four types of data: objective data (test score, submission time, and learning duration) is directly warehoused; Behavior data (click, stay, discussion posting) is stored after preprocessing (de duplication, exception filtering); Teachers' scores (experimental reports, project defense) are entered through the rubric, and each item has a clear scoring standard; The enterprise tutor's score is entered in a separate channel, and is weighted with the teacher's score to enter the overall evaluation.

Step 3: Fusion computing. According to the weight distribution in Section 4.3, the system calculates the comprehensive scores of each student periodically (weekly/module/final). At the same time, the learning analysis technology is used to generate a multi-dimensional competency profile:

Knowledge radar chart: show the mastery of six modules (highlight weak modules); Process line chart: show the change trend of classroom participation, submission timeliness and other indicators over time; Comparison heat chart: compare the individual student with the class average and the benchmark required by the enterprise.

Step 4: Feedback intervention. The system triggers personalized intervention according to the portrait results: real-time warning: when students have not submitted their homework for three consecutive times or the knowledge point test is less than 60 points, they will automatically send reminders to the teacher and student ends. Resource push: for weak knowledge points, push corresponding micro videos (such as "Paradigm Decomposition Teaching"), advanced exercises, and excellent open source project links of this knowledge point on GitHub. Teacher intervention: for the students who are in the bottom 10% of the comprehensive ranking, the system suggests that teachers arrange one-on-one Q&A or include them in the study group for key assistance.

Step 5: Continue iteration. At the end of each term, the teaching team conducts an overall analysis of the evaluation data: calculate the reliability (Cronbach's α) and validity (factor analysis) of each indicator, and eliminate the low discrimination indicators; Adjust the proportion or scoring standard of industry education integration indicators according to enterprise feedback; Update the "weak point mark" in the knowledge map (based on the high-frequency error data of previous students), and optimize the granularity of module division.

Through the above closed loop, the evaluation system itself also evolves, always in line with curriculum development and student needs.

5. Teaching Practice and Effect Verification

5.1. Practice Objects

This research takes the 2024 undergraduate students of the School of Computer and Software Engineering of Zhengzhou Sias University as the research object, involving computer science and technology and software engineering. A total of four natural classes were selected, two of which were set up as experimental classes (using the modular knowledge map enabling multi-source data fusion evaluation system built in this study), and the other two classes were combined into a control class (following the traditional evaluation model of "30% peacetime performance+70% final written examination").

The specific grouping is as follows: the experimental class includes Class 1 of Computer Science and Technology (62 students, 38 boys and 24 girls, with an average score of 78.3 in the early professional courses) and Class 1 of Software (60 students, 41 boys and 19 girls, with an average score of 77.9 in the early professional courses); The control class is a combination of Class 2 in Planning and Class 2 in Software (120 in mixed majors, 79 boys and 41 girls, with an average score of 78.1 in early professional courses).

The total number of effective samples was 242 (122 in the

experimental class and 120 in the control class). There is no significant difference among the three groups in terms of gender composition and average scores of early professional courses (one-way ANOVA, $F=0.32$, $p>0.05$), indicating that the grouping is balanced and comparative experiments can be conducted.

5.2. Practice Process

The practice cycle is the first semester of the 2024-2025 academic year, with 16 teaching weeks in total. It is divided into the following five stages:

Phase I: platform deployment and map binding (Week 1-2). Import the constructed knowledge map of six modules (more than 150 nodes, more than 200 relational edges) into the customized map module of Superstar Fanya Platform in JSON format. The teacher team bound all 29 evaluation tasks (including 7 chapter tests, 12 experimental reports, 4 PBL milestone reviews, 2 business tutor defense, and 4 classroom random tests) with relevant nodes in the knowledge map one by one. For example, in the "Third Experiment: Complex SQL Query" binding module 3, there are 8 nodes, including "Multiple Table Connection", "Sub Query", and "Aggregate Function". After the binding is completed, the platform will automatically generate a "list of covered knowledge points" for each task to ensure that the evaluation content and knowledge objectives are free of omission and redundancy.

The second stage: implementation of the teaching model (3-14 weeks). The trinity teaching mode of "knowledge map guidance+PBL project driven+multi-source data evaluation" is adopted. Before class: students watch module guidance videos (5-8 minutes for each module) on the Superstar platform, and complete the "pre-test" (5-10 objective questions) based on the knowledge map. The system automatically generates personalized preview suggestions according to the pre-test results. In class: the teacher explains the relationship between core knowledge points (such as "Why SQL optimization depends on indexes and execution plans") in combination with the visual presentation of the knowledge map, interspersed with in class quizzes and group discussions. The last 5 minutes of each class will be "Atlas Review". Students can review the content of the class by clicking the map node. After class: students complete the task by stages with the PBL project as the main line. This semester's PBL project is "Campus Card Consumption Analysis System Database Design", which is divided into five milestones: demand analysis (ER diagram), logical design (relational mode and paradigm), physical design (index and partition), SQL implementation (10 complex queries), and performance optimization (execution plan analysis). Each milestone has a clear scoring scale. Integration of production and teaching: in the 8th week, engineers from the cooperative enterprise (Dingli) will be invited to comment on and score the students' medium-term project plans through online meetings; In the final defense of the 16th week, the enterprise tutor and the teacher jointly form a review group, each accounting for 50% of the weight.

The third stage: multi-source data acquisition (whole process). Automatic record of the platform: learning behavior log (CSV is exported every day, including video viewing progress, test scores, forum posts, node click sequences), collecting about 87000 records in total. Teacher scoring entry: experimental reports (12 times), PBL milestones (5 times), classroom performance (once a week, 16 times in total), using a unified electronic scoring table (gauge), with 1-5 points for

each item. Enterprise tutor score: mid-term scheme score (hundred point system) and final defense score (hundred point system), which are collected by questionnaire stars and exported. Student self-evaluation and mutual evaluation: each group will conduct two mutual evaluations (the 8th and 16th week), scoring team collaboration, task commitment and other dimensions. All data are eliminated by Python script for outliers (such as deleting records with learning duration<1 minute) and filled with missing values (mean filling), finally forming a data matrix of 242 rows $\times$ 47 columns.

The fourth stage: fusion calculation and feedback (dynamic, once every two weeks). According to the weight in Section 4.3 (knowledge 30%, process 25%, practice 30%, literacy 15%), the system calculates students' real-time comprehensive scores every two weeks, and generates the overall class radar chart on the teacher side. For students whose comprehensive scores are less than 60 points or the mastery rate of a single module is less than 50%, the system will automatically mark and push reinforcement resources (such as designated micro lesson videos, exercises, and group coaching appointment links). In this semester, 347 personalized learning suggestions were put forward, and the teacher gave one-on-one guidance to 15 students according to the system prompts.

The fifth stage: final comprehensive evaluation and data analysis (the 15th to 16th weeks). The final written test (full score 100, the same test paper is used for the experimental class and the control class) is organized in a unified way, and the test content covers all six modules, including multiple choice questions (30%), blank filling questions (20%), short answer questions (20%), and comprehensive design questions (30%).

The experimental class additionally submitted complete PBL project source code and documents, while the control class completed the regular course design report (no enterprise scoring link). A questionnaire survey was carried out, and 238 valid questionnaires were recovered (the recovery rate was 98.3%), including learning satisfaction, cognitive load, evaluation fairness and other dimensions

(Likert's five level scale).SPSS 26.0 was used for independent sample t-test, Cohen's d calculation and chi square test.

5.3. Effect Analysis

This chapter mainly analyzes the effects of the experimental objects and the final experimental results obtained in the experimental process, and draws conclusions from the following four aspects.

The first is the level of students' knowledge structure, which is expressed by the mastery rate of knowledge points, the correct rate of knowledge related understanding, and the ability to draw frame maps. As shown in Table 4, the mastery rate of knowledge points: the average score of subjects involving basic concepts and principles (45 points in total) in the final written examination of the experimental class is 38.9 points, and the mastery rate is 86.4% (38.9/45); The average score of similar topics in the control class was 33.1 points, and the mastery rate was 73.6%. The difference between the two groups was significant ($t=5.32$, $p<0.001$, Cohen's $d=0.68$, medium effect). The correct rate of knowledge related understanding: in the 10 blank filling questions of "Fill in the predecessor/successor of knowledge points", the experimental class answered 7.8 correctly on average, with a correct rate of 78.0%; The average correct answer of the control class was 5.9, with an accuracy rate of 59.0%. Compared with the control class, the experimental class increased by 32.2% ($((7.8-5.9)/5.9 \approx 0.322)$, with a very significant difference ($t=6.14$, $p<0.001$). Framework drawing ability: At the end of the semester, students are required to draw the framework of the knowledge system of Database Principles by hand (not included in the total score). The frame diagram submitted by the experimental class contains 42.3 nodes and 58.6 relationship edges on average, while the control class has 28.7 nodes and 31.2 relationship edges on average. From the perspective of structural gestalt, 79.5% of the students in the experimental class were able to fully present the six modules and their progressive relationships, while only 41.7% of the students in the control class.

Table 4. Comparison of knowledge mastery between experimental class and control class

Index	Experimental class	Control class	Increase range	Significance
Mastery rate of knowledge points	86.4%	73.6%	12.8%	* $p<0.001$
Accuracy of knowledge association	78.0%	59.0%	32.2%	* $p<0.001$
Framework drawing ability	79.5%	41.7%	37.8%	* $p<0.001$

The results showed that the experimental class was significantly better than the control class in terms of knowledge structure, relevance understanding and system construction, which showed that the modular knowledge map effectively helped students establish global cognition.

The second is engineering practice ability, which is shown in Figure 3 from the following aspects. PBL project completion: the average score rate of the five milestones of the PBL project of the experimental class is 92.1% for demand analysis, 89.7% for ER design, 88.4% for logic design, 93.5% for SQL implementation, 86.2% for performance optimization, and 91.9% for the overall completion (the average of the five stages). The control class only completed the regular course design report (no milestone phased assessment), and the report quality score (hundred point system) averaged 74.3 points, significantly lower than the experimental class's PBL comprehensive score of 87.6 points ($t=8.23$, $p<0.001$). The excellent rate of enterprise tutors' score: in the overall score of the enterprise tutors of the

experimental class on the project scheme, more than 85 points (excellent) accounted for 82.8% (101/122), of which 4 groups (3.3%) obtained more than 95 points ("enterprise level excellent"); There is no enterprise tutor evaluation link in the control class. SQL query optimization ability: In the 45 minute "Query Optimization Experiment", the experimental class students completed three typical slow query optimization tasks, which took an average of 28.9 minutes, and the execution efficiency after optimization increased by 62.3% ($((\text{original time consumption} - \text{time consumption after optimization})/\text{original time consumption})$). The control class did not set up the experiment alone, but there was a SQL optimization question (15 points) in the final comprehensive design question. The average score of the experimental class was 12.7 points, and that of the control class was 9.4 points ($t=4.95$, $p<0.01$). Code normalization scoring: The enterprise tutor scored the project code of the experimental class according to Huawei OpenGauss coding specification (10 checkpoints), with a full score of 100 points and an average

score of 84.6 points. Among them, the score rate of "annotation clarity" is 86%, and the score rate of "indentation and naming rules" is 91%. This shows that students' engineering literacy has been effectively trained.

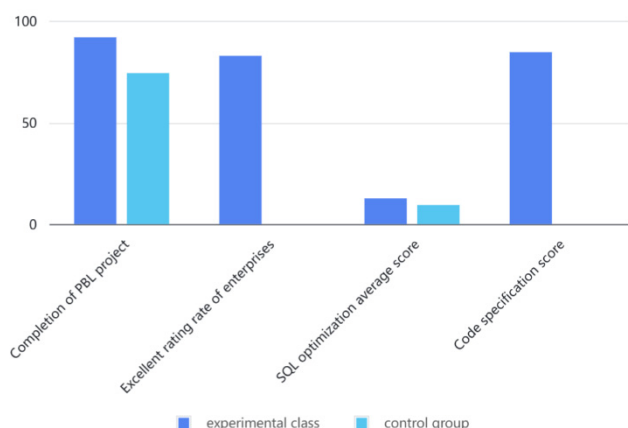


Fig 3. Comparison Diagram of Engineering Practice Capability

Third, learning depth and autonomy, as shown in Table 5, are mainly shown in the following aspects. Access depth of knowledge map: the platform log shows that the students in the experimental class visited the knowledge map nodes 183.7 times on average during the course, of which 52.8% of the behaviors whose access depth (continuous clicking on

related nodes) exceeded three levels (that is, more than half of the access sequences involved deep reasoning), while the control class did not use the knowledge map, and there was no comparative data. However, compared with the previous historical data of traditional teaching classes (2023, no atlas), the proportion of in-depth interviews increased by about 40 percentage points. Independent inquiry behavior: the average number of times that the experimental class students actively post questions in the "technical discussion area" of the forum is 2.7 times per person, and the number of times that they actively answer others' questions is 1.9 times per person; The corresponding numbers of the control class were 1.1 times and 0.6 times respectively. The experimental class's active questioning involved high-level cognitive problems such as "comparison of optimization methods" and "exception scene processing", accounting for 61.3%, while the control class mostly involved low-level problems such as "syntax errors" and "installation configuration". Learning duration distribution: the average weekly platform learning duration of the experimental class is 4.2 hours, of which 73.8% are more than 3 hours; The average person in the control class was 2.9 hours per week, and only 45.0% of the students had more than 3 hours per week. More importantly, the study duration of the experimental class in the non mandatory period (weekends, after 21:00 in the evening) accounted for 34%, while that of the control class was only 19%, reflecting a stronger intrinsic motivation.

Table 5. Comparison of learning depth and autonomous learning behavior

Index	Experimental class	Control class	Difference
Map access depth>3 layers	52.8%	—	Significant improvement
Average weekly autonomous learning duration	4.2 hours	2.9 hours	Increase by 45%
Number of active questions/answers	2.7/1.9	1.1/0.6	Significant improvement
Proportion of learning in non mandatory period	34%	19%	Increase by 15 percentage points

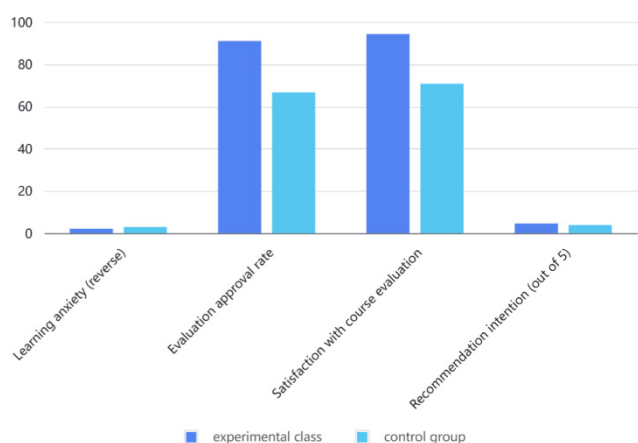


Fig 4. Comparison of learning experience and satisfaction

Finally, learning experience and satisfaction, as shown in Figure 4, include learning anxiety: the questionnaire uses the "Learning Anxiety Scale" (5 questions, 5 grades), the average anxiety of the experimental class is 2.1 points (1=very low anxiety, 5=very high), and the control class is 2.9 points, down 27.6%. Especially in the item of "worrying about failing the exam", the average value of the experimental class is 1.8, and that of the control class is 2.7. Evaluation of sense of fairness: the question "I think the course scoring method is objective and fair", the proportion of agreement and very agreement in the experimental class is 91.0%, and that in the control class is 66.7%. A student left a message: "Each assignment and project can see the score of the corresponding

knowledge points. Where there is weakness, there is no sense of being" determined by one exam ". Course satisfaction: 94.3% (115/122) of the experimental class and 70.8% (85/120) of the control class are satisfied with the "course evaluation model ". The satisfaction with the "overall curriculum" of the experimental class was 91.8%, and that of the control class was 76.2%. Willingness to recommend: "I am willing to recommend this course to younger students". The average willingness to recommend in the experimental class is 4.6/5.0, while that in the control class is 3.9/5.0.

5.4. Teacher Application Feedback

At the end of the semester, semi-structured interviews were conducted with three teachers (one associate professor and two lecturers) who undertook the teaching tasks of the experimental class and two teaching assistants who participated in data management, and the following feedback was obtained in combination with the system use log. 1、Assessment workload: all three teachers believe that after the adoption of this system, data collection and analysis are automatically completed by the platform, and teachers only need to score according to the scale (equivalent to the previous workload). However, at the end of the period, it is no longer necessary to manually summarize the usual scores, conversion ratio, etc., and the overall evaluation of routine work time is reduced by about 35% (the average shift time is reduced from 8 hours to 5.2 hours). 2、Teaching assistant feedback: the data cleaning script reduces the weekly data processing time from 2 hours to 0.5 hours. Evaluation

accuracy improvement: teachers believe that because each evaluation task is bound to knowledge nodes, students' weak points can be accurately located (such as the "transaction isolation level" node highlighted in red in the map). Teachers have made up five concentrated lessons for common problems in the class this semester, and the mastery rate of this node has increased by 22% on average after each make-up lesson. In the past, it was difficult to judge with such accuracy based on experience. 3、Convenience of learning situation analysis: teachers can view each student's knowledge radar chart and process trend line in real time through the platform visualization panel. Teacher A said: "Previously, I could only see the total score, but now I can see which module the student card is in. For example, the average mastery rate of the class at the 'paradigm' node is only 63%, so I added a special exercise class." Teacher B feedback: "The score of the enterprise mentor let me know that there is still a gap between the students and their real positions, and I plan to add version control training next semester." 4. Technology mastery: There are 5 teachers and teaching assistants, and after two trainings (2 hours each time), 4 (80%) can independently complete core operations such as knowledge map node binding, evaluation task configuration, and viewing information reports. An older lecturer needed technical support, but said that it was "simpler than expected, mainly click and select operations, without writing code". The preset templates in the system greatly reduce the threshold.

Based on the above feedback, the evaluation system has not only improved the scientificity and efficiency of the evaluation, but also been positively recognized by the teacher team.

6. Research Conclusion and Prospects

6.1. Research Conclusion

This research has constructed a multi-source data fusion evaluation system empowered by the modular knowledge map, and has been verified by 242 students' teaching practice. The conclusions are as follows: The modular knowledge map is divided into six modules (more than 150 nodes, more than 200 relational edges) to provide accurate knowledge anchor points; The multi-source data fusion evaluation system integrates four types of data, with the weights of knowledge 30%, process 25%, practice 30%, and literacy 15%, forming a closed-loop mechanism; The experimental class was significantly better than the control class in terms of knowledge structure level (86.4% vs 73.6%), engineering practice ability (PBL completion rate 91.9%), learning autonomy and satisfaction (94.3% vs 70.8%). The workload of teachers was reduced by about 35%, and 80% of teachers could quickly master technical operations. This system provides a replicable practice paradigm for the curriculum evaluation reform of private colleges and universities.

6.2. Research Deficiencies and Future Prospects

This study also has some limitations: the sample is limited to one private university for one semester, the depth of enterprise evaluation is insufficient, the degree of intelligent data fusion needs to be improved, privacy and ethical issues need further attention, and the reuse of knowledge maps across courses needs the support of standardized tools. In the future, it will expand the sample range, deepen the integration mechanism of industry and education, introduce machine learning to realize dynamic weight optimization, improve the data ethics framework, and build a low code map co construction platform to promote the large-scale application of the evaluation system in more scenarios.

Acknowledgments

The 2025 Educational Reform Fund Project of Sias University (Project Number: 2025JGYB43).

References

- [1] Ma, R., Zhao, L., Li, Z., et al. (2025). *Application of digital teaching materials based on competency mapping*. Alliance for Innovation and Development of Digital Teaching Materials in Higher Education.
- [2] Lu, Y., Fan, Q., & Zhou, J. (2025). Research and practice of "AI+ knowledge map" enabling the construction of golden courses in higher vocational colleges -- Taking the course "openGauss database technology" as an example. *China Information Technology Education*, (9).
- [3] Wang, W., & Yang, F. (2025, October 11). "XD Smart Evaluation" - Exploration and Practice of Digital Smart Enabling Student Evaluation. Xi'an University of Electronic Science and Technology News Network.
- [4] Xie, R., Zhang, S., & Hu, Y. (2025). AI driven reform of classroom teaching quality evaluation and innovation of intelligent teaching practice. *Electronic Journal of Henan Education and Research Network*.
- [5] Yi, W., Huang, X., Kuzmin, S., et al. (2025). Seekg: Sentiment Analysis for E-Learning Evaluation Incorporating Knowledge Graphs. *Education and Information Technologies*, 30(12), 16291–16320. <https://doi.org/10.1007/s10639-025-13456-1>.
- [6] Shi, H., Pan, G., & Fan, H. (2025). Research on intelligent reconstruction and mechanism optimization of MOOC course teaching evaluation system. *Journal of Water Resources and Building Engineering*, (4), 240–246.
- [7] Tang, F., Chen, M., Xu, J., et al. (2025). Instructional Design and Practice of Specialized Courses Based on Knowledge Graphs—Using the Fundamentals of Electrical Engineering as a Case Study. *Frontiers of Digital Education*, 2(1), 7. <https://doi.org/10.1007/s44366-025-0050-x>Higher Education Press.