



# Research on Constructing Student Digital Twins Driven by Digital-Intelligent Integration Based on Disciplinary Knowledge Graphs

Tao Zhao\*. Dan Wu, Jun Cao

Chongqing Vocational college of Culture and Arts, Banan, Chongqing, 401320, China

\*corresponding author: taozhaocqcaedu@163.com

**Abstract.** Against the backdrop of global educational digital transformation, traditional higher education struggles to meet diverse learning needs and talent market demands. This study proposes constructing a digital-intelligent integration-driven Student Digital Twin (DTS) based on a six-layer disciplinary knowledge graph. The disciplinary knowledge graph provides a semantic core and knowledge foundation, while the DTS integrates real-time academic data, behavioral signals, and competency levels. Through three dynamic coupling cycles—state perception-semantic mapping, intelligent analysis-path reasoning, and intervention feedback-graph evolution—deep integration between the static disciplinary knowledge graph and the dynamic DTS is achieved. This paradigm transforms unified knowledge transmission into personalized competency development and achieves innovations in three key domains of education: Firstly, in the teaching domain, it enables the shift from experience-driven to evidence-driven precision teaching; Secondly, in the learning domain, it facilitates the transition from passive knowledge reception to active personalized growth; Thirdly, in the evaluation domain, it realizes the transformation from outcome-oriented assessment to process-oriented value-added competency certification. These practical innovations in the three domains provide a feasible technical architecture and practical path for the competency-oriented transformation of higher education, reshaping a new talent training ecosystem in the digital era.

**Keywords:** Student Digital Twin, Disciplinary Knowledge Graphs, Digital-Intelligent Integration

## 1 Introduction

This study takes the construction of disciplinary knowledge graphs as the foundation and positions the digital-intelligent driven Student Digital Twin (DTS) as a core element of the future learning architecture for college students. There is a growing recognition that building a personalized, flexible, and competency-oriented learning ecosystem is crucial for helping students stand out in the fiercely competitive talent market <sup>[1]</sup>. A student digital twin is defined as a continuously evolving digital representation of a student, integrating real-time academic data, behavioral signals, competency levels, and career preferences <sup>[2]</sup>. It serves dual functions as both a diagnostic

© The Author(s) 2026

S. Garcia-Esteban et al. (eds.), *Proceedings of the 2026 5th International Conference on Social Sciences and Humanities and Arts (SSHA 2026)*, Advances in Social Science, Education and Humanities Research 1014, [https://doi.org/10.2991/978-2-38476-577-5\\_60](https://doi.org/10.2991/978-2-38476-577-5_60)

tool and a prescriptive engine, enabling two-way interaction: on the one hand, AI algorithms enrich the content of the digital twin based on learner data; on the other hand, the student digital twin provides a basis for generating and adjusting personalized, competency-oriented learning paths.

## **2 The "Six-Layer" Architecture System of Disciplinary Knowledge Graphs**

The six-layer architecture of this disciplinary knowledge graph is a typical end-to-end "data-knowledge-application" system, As shown in Figure 1 below with the professional functions of each layer analyzed as follows:

### **2.1 Heterogeneous Data Layer (Data Input Layer)**

This layer serves as the data source foundation for the entire construction of disciplinary knowledge graphs. It aggregates various types of raw data in the disciplinary domain, which can be classified into three categories. Structured data: Regularized data with fixed row-and-column structures and unified data types, such as disciplinary databases, course schedules, and academic grade sheets. Semi-structured data: Data with flexible and extensible structures that do not conform to rigid tabular formats. Unstructured data: Data without fixed structures or predefined fields, mainly in the form of multimedia content or free texts. The role of this data layer is to provide comprehensive raw materials for subsequent knowledge processing, and to address the problems of scattered disciplinary data and inconsistent data formats.

### **2.2 Knowledge Acquisition Layer (Information Extraction Layer)**

This layer acts as the conversion gateway that transforms data into knowledge. Its core task is to extract disciplinary knowledge elements from heterogeneous data based on Natural Language Processing (NLP) technology. The specific steps are as follows: firstly, conducting entity recognition to extract core disciplinary concepts; secondly, performing entity classification to categorize the identified entities; finally, carrying out relation extraction by mining the associations between entities. The primary function of this layer is to convert unstructured or semi-structured data into preliminary knowledge units in the form of "entity-relationship", which serves as a fundamental link in the construction of disciplinary knowledge graphs.

### **2.3 Knowledge Fusion Layer (Knowledge Alignment Layer)**

This layer functions as the knowledge standardization and integration layer. Its core task is to resolve the issues of redundancy, conflict and heterogeneity in knowledge derived from multi-source data, so as to achieve knowledge unification. The implementation steps include: firstly, implementing schema fusion to unify the top-level framework of disciplinary knowledge; secondly, conducting entity alignment to

merge different expressions referring to the same entity; finally, performing entity linking to associate the extracted entities with the existing knowledge systems. The main role of this layer is to eliminate knowledge ambiguity, form a logically consistent and structurally unified set of disciplinary knowledge, and enhance the standardization of knowledge.

#### **2.4 Knowledge Storage Layer (Knowledge Carrier Layer)**

This layer serves as the structured knowledge storage and quality control layer. Its core task is to store the fused knowledge in a graph structure (with entities as nodes and relationships as edges) and ensure the quality of knowledge. The key operations are as follows: adopting graph databases to store the knowledge graphs in the form of "entity-relationship-attribute" triples; conducting knowledge graph quality assessment by screening erroneous or low-quality knowledge through consistency verification and completeness detection. The function of this layer is to enable efficient storage, retrieval and management of knowledge, providing a reliable knowledge carrier for subsequent reasoning.

#### **2.5 Knowledge Reasoning Layer (Knowledge Enhancement Layer)**

This layer acts as the knowledge expansion and decision-making layer. Its core task is to derive new content and optimize the knowledge system based on existing knowledge. The specific processes are as follows: firstly, implementing knowledge completion by filling the missing relationships in the knowledge graph; secondly, conducting knowledge denoising by identifying and eliminating contradictory or erroneous knowledge; finally, performing logical reasoning based on the disciplinary knowledge graph. The role of this layer is to enrich the content dimensions of the knowledge graph and improve the depth and usability of knowledge.

#### **2.6 Knowledge Empowerment Layer (Application Output Layer)**

This layer serves as the value realization layer of the disciplinary knowledge graph. Its core task is to transform the capabilities of the knowledge graph into application services for disciplinary scenarios, mainly including four aspects: firstly, semantic search for accurate information retrieval based on knowledge associations; secondly, intelligent recommendation for personalized delivery of disciplinary resources; thirdly, enabling intelligent question-answering for disciplinary knowledge; finally, explanation generation to provide interpretable accounts for reasoning results. The function of this layer is to connect scenarios such as teaching, learning and scientific research, and transform the disciplinary knowledge graph into practical tools<sup>[3]</sup>.

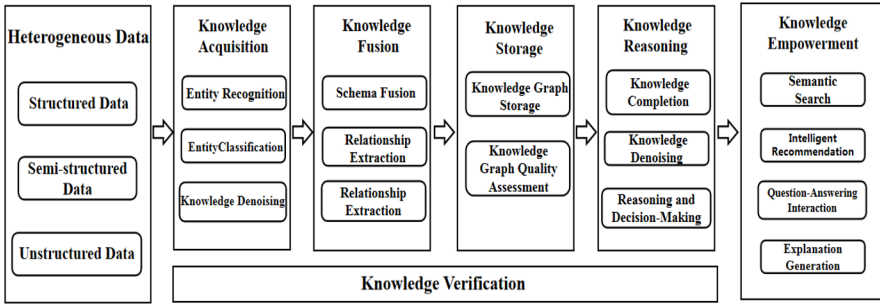


Fig. 1. The Six-Layer Architecture of the Disciplinary Knowledge Graph.

### 3 System Theoretical Paradigm: Constructing Digital-Intelligent Integrated Digital Twins Based on Disciplinary Knowledge Graphs

The six-layer "data-knowledge-application" architecture of disciplinary knowledge graphs provides a semantic core and knowledge foundation for the digital-intelligent integrated Student Digital Twin (DTS) in higher education. The coupling of the two constitutes a key paradigmatic innovation for realizing personalized, competency-oriented learning.

The heterogeneous data layer, knowledge acquisition layer, and knowledge fusion layer of the knowledge graph convert structured, semi-structured, and unstructured data in the disciplinary field into standardized "entity-relationship-attribute" triples through technologies such as entity recognition and Schema alignment. This provides semantic anchors for disciplinary knowledge for the digital twin, defining a reference frame for "what constitutes disciplinary knowledge" and a logical framework for the semantic mapping of scattered learning data<sup>[4]</sup>. Meanwhile, the graph database storage and logical reasoning capabilities of the knowledge storage and reasoning layers serve as the core basis for the digital twin to conduct competency gap analysis: leveraging the association rules of the knowledge graph, the digital twin can accurately identify students' knowledge breakpoints and realize the quantitative conversion from "knowledge state" to "competency dimension".

### 4 System Implementation Mechanism: Dynamic Coupling between Knowledge Graphs and Digital Twins

To realize the aforementioned paradigm, it is necessary to design a core mechanism that enables in-depth, real-time coupling between static disciplinary knowledge graphs and dynamically evolving student digital twins. This mechanism consists of the following three key cycles:

#### 4.1 State Perception and Semantic Mapping Cycle

This cycle is the foundation of data-driven operations. The student digital twin continuously collects students' academic performance, online behaviors, interaction patterns, and even emotional signals by integrating multi-source data pipelines such as Learning Management Systems (LMS), intelligent learning tools, and IoT sensing devices. These raw data streams are semantically mapped to specific nodes and relationships in the disciplinary knowledge graph through entity recognition and relationship extraction technologies in the knowledge acquisition layer. For example, multiple failed attempts at a calculus exercise are no longer merely a statistical value of "number of errors" but are mapped to the entity node "chain rule" in the knowledge graph, associated with its prerequisite knowledge point "composite function" and subsequent knowledge point "implicit function differentiation", thereby forming a semantically meaningful "knowledge mastery state".

#### 4.2 Intelligent Analysis and Path Reasoning Cycle

This cycle is the core of intelligent decision-making. Based on the student's knowledge state after semantic mapping, the knowledge reasoning layer activates its logical engine. Utilizing predefined rules and machine learning models in the knowledge graph, it conducts two key analyses: Firstly, competency diagnosis and gap identification: Through algorithms, it analyzes the gap between the student's mastered knowledge points, target competencies, and the required knowledge and competency graph, accurately locating weak links and breakpoints in the competency chain. Secondly, personalized path generation: Combining the diagnostic results with metadata such as students' career preferences and learning styles, the reasoning layer optimizes and generates multiple potential learning path sequences within the extensive knowledge graph. This is not a simple linear recommendation but dynamic programming based on neural networks, capable of recommending non-traditional paths such as "detour reinforcement" or "interdisciplinary transfer".

#### 4.3 Intervention Feedback and Graph Evolution Cycle

This cycle ensures the adaptability and vitality of the system. The digital twin implements interventions by delivering the generated personalized paths through recommendation engines, virtual learning companions, or teacher dashboards. Students' feedback on interventions is captured by the perception layer as new data streams.<sup>[5]</sup> This feedback is not only used to update the state of the digital twin but, more importantly, feeds back to the disciplinary knowledge graph itself through online learning mechanisms, specifically through the following two paths: Firstly, dynamic adjustment of knowledge relationship weights: If a large number of students frequently struggle with the path "knowledge point A to B", the system may automatically lower the default weight of this path or trigger an alert prompting curriculum designers to review the teaching sequence. Secondly, discovery of implicit knowledge relationships: By analyzing the commonalities of successful learning paths among a large

number of students, machine learning models may identify effective knowledge associations not predefined in the graph, thereby expanding and optimizing the connotation of the knowledge graph. This transforms it from an expert-constructed static model into a "living" knowledge ecosystem that aggregates collective wisdom and continuously learns.

## **5 Research Conclusions**

This study proposes the construction of student digital twins driven by digital-intelligent integration based on disciplinary knowledge graphs, covering the construction of the "six-layer" architecture system of disciplinary knowledge graphs, the theoretical paradigm of system construction, and the implementation mechanism of system construction, thereby realizing the dynamic coupling between knowledge graphs and digital twins. The construction of this coupling mechanism has completely transformed traditional teaching in the fields of teaching, learning, and regulation and evaluation.

### **5.1 Evolution of the Teaching Field: From Experience-Driven to Evidence-Based Precision Teaching**

In the teaching field, it promotes the shift from experience-driven to evidence-based precision teaching. Teachers can design differentiated teaching based on the group portraits of class digital twins, dynamically adjust strategies during classes based on real-time feedback, and implement large-scale personalized tutoring through automated grouping and resource push after classes. This enables "teaching students in accordance with their aptitude" to be effectively realized in a data-driven and operable manner.

### **5.2 Subversion of the Learning Field: From Passive Acceptance to Active Construction of Personalized Growth**

In the learning field, it facilitates students' transition from passive acceptance to active construction of personalized growth. Through visual "academic mirrors", students can clearly understand their own competency structures and obtain personalized learning navigation dynamically generated based on real-time states. This process not only adapts to individual paths but also deepens students' understanding of the inherent connections between knowledge through the system's explanation function, effectively cultivating their metacognitive and self-directed learning abilities.

### **5.3 Focus of the Regulation and Evaluation Field: From Outcome-Oriented Evaluation to Process-Oriented Value-Added Competency Certification**

In the field of regulation and evaluation, it realizes the transformation from outcome-oriented evaluation to process-oriented value-added competency certification. Focusing on students' academic progress and value-added evaluation during the learning process, the system generates a "digital competency profile" framed by the knowledge graph, which goes beyond traditional transcripts to three-dimensionally present students' interdisciplinary competency networks. Meanwhile, aggregated data provides precise insights for curriculum system optimization and industry-education integration, supporting the real-time iteration of school talent training programs. It also closely monitors the labor market's demand for knowledge and skills with keen insight, enabling schools to targeted enhance students' knowledge and skill development through macro or micro regulation.

Finally, the framework proposed in this study for constructing student digital twins driven by digital-intelligent integration based on disciplinary knowledge graphs responds to the core demands of personalized and competency-based education in the digital transformation of education. It elevates disciplinary knowledge graphs from static "knowledge maps" to dynamic "semantic operating systems" and evolves student digital twins from simple data dashboards to "intelligent learning agents" with diagnostic and prescriptive capabilities. The in-depth integration of the two provides an operable technical path and theoretical paradigm for realizing the fundamental transformation from "unified teaching" to "personalized growth" and from "knowledge transmission" to "competency construction". This is not only an upgrade of learning technologies but also a systematic reconstruction of the educational logic in higher education, which is expected to reshape the talent training ecosystem in the digital era, enabling each student's development path to possess both disciplinary depth and individual breadth.

## **Acknowledgement**

This work was sponsored in part by the 2025 Chongqing (China) Education Science Planning Project "Research on the Mechanism and Path of AIGC Enhancing University Teachers' Teaching Competence Based on the TPACK Model" (Project No.: S25YG3360002) and the 2024 Science and Technology Research Project of Chongqing Municipal Education Commission (China) "Research on Constructing Student Digital Twins Based on Knowledge Graph Driven by Digital-Intelligent Integration" (Project No.: KJQN202406202).

## References

1. Sáez-López, J.-M.; Domínguez-Garrido, M.-C.; Medina-Domínguez, M.-d.-C.; Monroy, F.; González-Fernández, R. The Competences from the Perception and Practice of University Students. *Soc. Sci.* 2021, 10, 34.
2. Igor Kabashkin. AI-Based Digital Twins of Students: A New Paradigm for Competency-Oriented Learning Transformation[J]. *Information*, 2025, 16, 846.
3. Sun Honglin, Hou Wu. Knowledge Graphs: From Theory to Practice [M]. Beijing, China: Tsinghua University Press, 2025.
4. MA Fulong. Subject Knowledge Graph: Connotation, Technique Architecture, Application and Development Trend[J]. *Software Guide*. 2024(23)3:213-216.
5. AI Xing, Zhang Yu. From Digital Portrait to Digital Twin: A New Exploration of the Construction of Digital Twin Learner Driven by the Integration of Big Data and Artificial Intelligence[J]. *Journal of Distance Education*. 2021(01):41-47.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

