Advances in Vocational and Technical Education (2026)
Clausius Scientific Press, Canada
DOI: 10.23977/avte.2026.080104
ISSN 2523-5834 Vol. 8 Num. 1

Evolutionary Trends and Structural Characteristics of Chinese Vocational Computer Curricula in the AI Era: A Semantic Network Analysis

Zhuohuang Zhang*, Li Li

School of Artificial Intelligence, Shenzhen Polytechnic University, Shenzhen, China
**Corresponding Author*

Keywords: Artificial Intelligence, Vocational Education, Curriculum Reform, Word2Vec, Semantic Network

Abstract: The rapid advancement of Artificial Intelligence (AI) has driven significant changes in vocational computer curricula in China. Unlike previous studies focused on theoretical models, this research uses a data-driven approach to evaluate the structural evolution of these reforms within the Chinese vocational education context. Utilizing the Word2Vec model and semantic network analysis, we examined 370 core academic papers published between 2018 and 2025. The empirical results reveal three critical structural characteristics. First, we identify a distinct structural stratification in curriculum integration, where application-layer courses demonstrate high semantic coupling with AI, whereas infrastructure-layer courses remain relatively isolated. Second, we observe a pragmatic turn in the temporal evolution of reform focus; post-2022, the discourse decisively shifted from macroscopic policy design to microscopic pedagogical practice. Third, we detect a pronounced asymmetry in stakeholder attention, characterized by a heavy concentration on technological tools and student outcomes, contrasted with the marginalization of teacher digital literacy. These findings highlight the need to strengthen foundational courses and enhance faculty capabilities for a more balanced educational structure.

1. Introduction

The integration of Artificial Intelligence (AI) into higher vocational education has transformed from an emerging trend to a structural necessity. Global consensus suggests that AI is fundamentally reshaping the skill requirements for the future workforce, requiring a fundamental shift in engineering education[1]. With the advent of generative AI technologies such as ChatGPT, the foundational logic of computer science education is undergoing a profound disruption, shifting focus from routine code generation to high-level system architecture and prompt engineering[2]. In response to these shifts, national policies in China have increasingly emphasized the digitalization of vocational education, mandating that institutions adapt their curricula to meet the evolving demands of the intelligent industry. Consequently, vocational colleges have launched extensive reforms, ranging from the introduction of "AI + X" interdisciplinary courses to the restructuring of traditional programming instruction[3].

However, the efficacy and depth of these reforms remain subjects of debate. Existing literature on this transition often relies on qualitative experience summaries or traditional bibliometric methods based on keyword frequency[4]. While these approaches identify broad trends, they often fail to capture the latent semantic associations and structural evolution within educational discourse. For instance, traditional Bag-of-Words models cannot distinguish whether a "reform" is merely a macroscopic slogan or a microscopic implementation of code practice[5]. Consequently, it remains unclear whether these reforms have impacted actual teaching practices or remain at the level of conceptual advocacy.

To address this gap, this study proposes a text mining framework based on the Word2Vec neural network model. By analyzing a corpus of 370 core papers from the China National Knowledge Infrastructure (CNKI), covering the period from 2018 to 2025, this research aims to provide a data-driven assessment of the reform trajectory. The study focuses on three dimensions: the penetration depth of AI across different curriculum layers, the temporal evolution of reform granularity, and the distribution of attention among key educational stakeholders. By rigorously examining these dimensions, this paper seeks to provide empirical evidence for policy-making in the next phase of vocational education reform.

2. Experimental Design

2.1. Data Acquisition and Preprocessing

The dataset for this study was derived from the China National Knowledge Infrastructure (CNKI) database, chosen for its authoritative coverage of Chinese policy discourse and educational research. To ensure the representativeness of the sample regarding the "AI + Vocational Education" domain, we employed a targeted search strategy combining "Artificial Intelligence" with "Curriculum Reform" within the subject areas of "Computer Science" and "Vocational Education." The time span was restricted to the period between January 1, 2018, and 2025, covering the complete cycle from the initial policy advocacy of AI education to the recent explosion of Generative AI.

A rigorous manual screening process was implemented to exclude non-substantive entries such as news reports, meeting notices, and broad policy interpretations, specifically prioritizing papers that presented empirical curriculum schemes or specific teaching cases. This filtering resulted in a final corpus of 370 high-quality academic papers, which includes CSSCI journals and core vocational education periodicals, ensuring that the analysis reflects the highest level of academic discourse in the field.

2.2. Semantic Network Construction

To overcome the limitations of traditional bibliometric methods, which often rely on frequency counts and fail to capture context, this study utilized the *Word2Vec* neural network model to map vocabulary into a high-dimensional vector space[6]. We utilized the Skip-gram architecture with a vector dimension of 300 and a window size of 5 to optimally capture the semantic proximity between technical terms in short text contexts. The processing pipeline involved three distinct stages. First, unstructured text data was extracted from PDF documents using the *pdfplumber* library to ensure complete retrieval of content, particularly preserving the structural layout of tables and diagrams where curriculum details are often located. Second, the text underwent Chinese word segmentation via the *jieba* toolkit, followed by the removal of stop words using a customized academic exclusion list to enhance the signal-to-noise ratio. Finally, a dynamic semantic dictionary was constructed using a Seed Expansion Algorithm, which calculated the cosine similarity between candidate terms and seed words to automatically identify domain-specific synonyms before human verification.

To ensure the domain validity of the analysis, we manually inspected the raw keyword clusters generated by the AI model (e.g., for technology stacks or curriculum names) and filtered out irrelevant noise and generic terms. This rigorous screening process allowed us to build highly refined semantic dictionaries that served as the precise basis for our subsequent quantitative experiments.

2.3. Experimental Design Framework

The research design encompasses three distinct experimental dimensions to systematically evaluate the depth and efficacy of the curriculum reform, forming a triangulated analytical framework. First, a co-occurrence matrix was constructed crossing "Computer Core Courses" with "AI Technology Stacks." This dimension aims to diagnose the structural integrity of the reform by visualizing the semantic distance between specific courses and AI technologies, determining whether the reform is a deep "ecological reconstruction" or merely a superficial "tool overlay." Second, a longitudinal analysis of reform granularity was conducted by establishing two opposing semantic sets: "Macro-Design" and "Micro-Practice." By tracking their normalized frequency annually from 2018 to 2025, a metric calculated relative to the total token count per year to mitigate biases from varying publication volumes, this dimension verifies the "Pragmatic Turn" hypothesis, particularly following the emergence of Generative AI in late 2022. Third, we modeled the subject attention distribution by quantifying the weighted attention assigned to "Technology," "Teachers," and "Students," aiming to detect potential structural imbalances in the digital transformation ecosystem.

Methodologically, these three dimensions constitute a cohesive, triangulated analytical framework designed to capture the multidimensional complexity of educational reform. The first dimension (Semantic Penetration) provides a static structural diagnosis, mapping the spatial disparities of AI integration across the curriculum system; the second dimension (Reform Granularity) introduces a dynamic temporal axis, revealing how the reform focus adapts to external technological shocks over time; and the third dimension (Subject Attention) offers an agent-based perspective, evaluating the resource allocation among the critical human stakeholders. By integrating these spatial, temporal, and human-centric metrics, the framework moves beyond traditional descriptive bibliometrics to construct an explanatory model of curriculum evolution. To ensure analytical precision, all quantitative indicators—such as cosine similarity for semantic distance and normalized frequency for temporal trends—were computed using standard Python scientific computing libraries, with strict data normalization protocols applied to mitigate biases arising from annual fluctuations in publication volume.

3. Results

3.1. Semantic Penetration

We present the "Curriculum-Technology" co-occurrence matrix in Figure 1. Our analysis reveals that application-layer courses, such as *Python Development* and *Program Design*, exhibit a high degree of integration with AI terminology. Specifically, the co-occurrence frequency of Python with Artificial Intelligence reaches 86, indicating a dominant trend of "Tool Substitution," where AI is primarily utilized to update programming languages. In contrast, infrastructure-layer courses that underpin computer science fundamentals, such as *Operating Systems* and *Computer Networks*, show comparatively weaker semantic association with AI terms. This disparity indicates a structural stratification, where reforms are predominantly skewed towards application-layer technologies, resulting in a marked deficit in the integration of AI concepts into foundational system curricula.

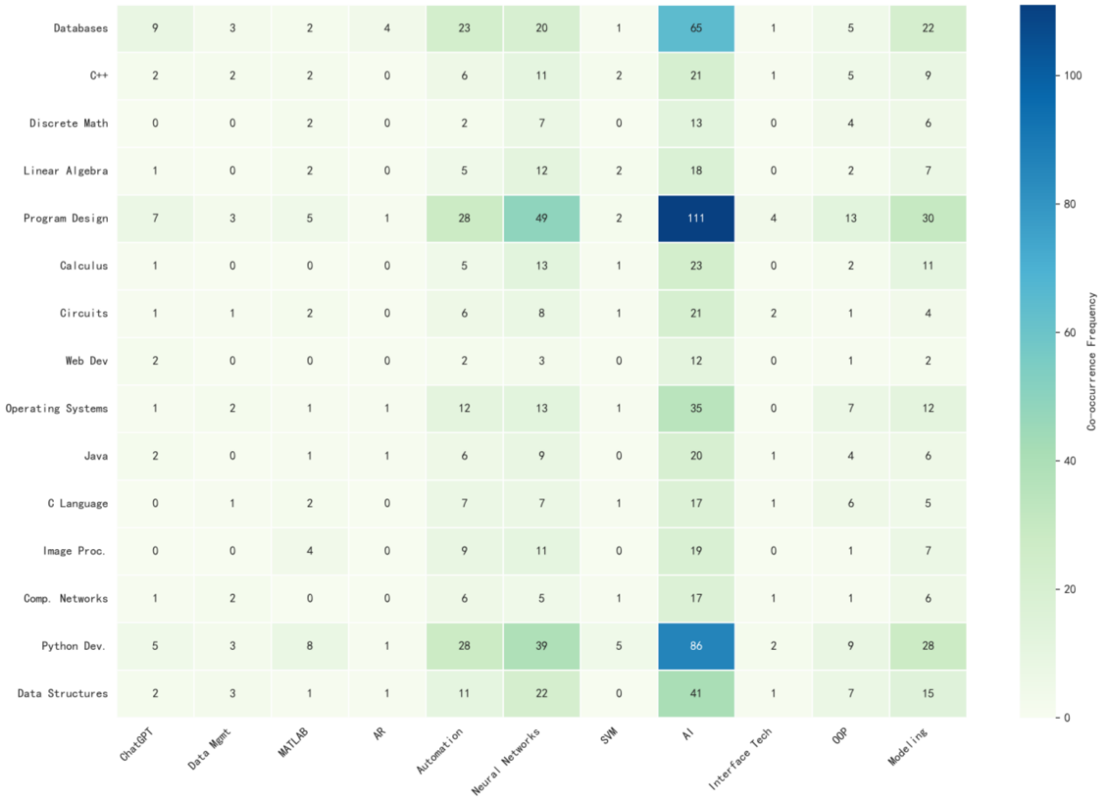

Figure 1. The semantic penetration heatmap of computer core courses and AI technologies.

3.2. Evolutionary Trajectory

To systematically examine the implementation depth of the reforms, we categorized semantic keywords into three granular levels: Macro Design (L1), encompassing systemic concepts such as "top-level design" and "talent training models"; Resource Building (L2), focusing on infrastructural elements like "digital platforms" and "online resources"; and Micro Practice (L3), centering on tangible pedagogical executions such as "code debugging" and "hands-on projects."

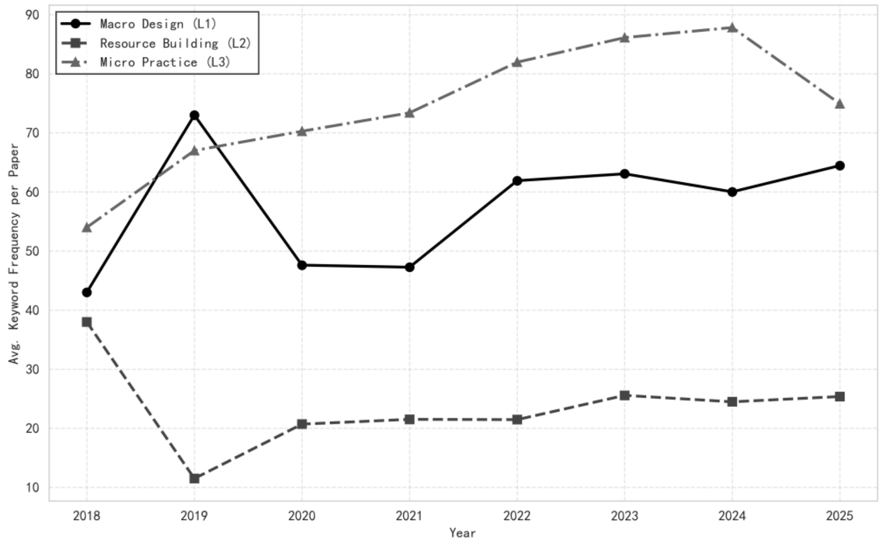

Figure 2. The evolutionary trend of reform granularity from 2018 to 2025.

A longitudinal analysis of these dimensions (Figure 2) reveals a distinctive evolutionary pattern characterized by the eventual dominance of micro-level practices. During the initial phase (2018–2022), the reform discourse exhibited a trend of "Entangled Fluctuation," where attention levels between macro-level planning and micro-level practice were relatively comparable. Notably, the Macro Design dimension experienced a significant spike in 2019, reaching a normalized frequency of 73.0 and temporarily surpassing Micro Practice (67.0), a fluctuation likely attributed to policy-level advocacy during the early stages of AI integration. However, a structural divergence became evident post-2022. The Micro Practice dimension subsequently established a robust and consistent upward trajectory, surging to a peak of 87.8 in 2024. In contrast, Macro Design stabilized in the range of 60.0 to 64.5, consistently trailing behind the micro-practice dimension in the later stages. Meanwhile, Resource Building remained relatively stable at a lower level (hovering between 20.0 and 25.0), suggesting that while digital resources are essential, the primary focus of the reform discourse has decisively shifted beyond mere resource accumulation towards concrete teaching implementation and specific pedagogical scenarios.

3.3. Attention Distribution

In the third dimension, we analyzed the weighted frequency of educational stakeholders (Figure 3), revealing a pronounced structural asymmetry within the current reform landscape. We find that the discourse is disproportionately concentrated on the endpoints of the educational process: "Student-Centric" attention scores the highest (15,220), followed closely by "Technology-Centric" attention (12,026). In sharp contrast, "Teacher-Centric" attention remains significantly lower (5,822). We interpret this imbalance as evidence of "Technological Determinism," where the reform narrative prioritizes the deployment of digital tools and the measurement of student outcomes, while comparatively marginalizing the adaptive agency of teachers. This finding highlights a critical vulnerability in the transformation strategy, as it underemphasizes the role of educators as the essential mediators between technological infrastructure and pedagogical effectiveness.

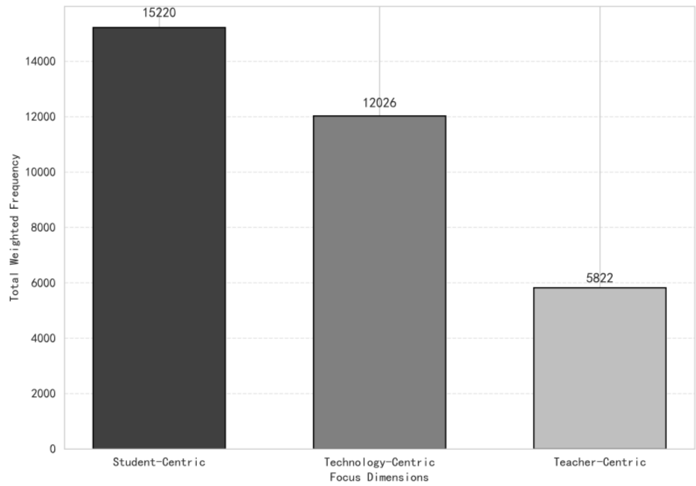

Figure 3. Comparison of subject attention focus: Technology vs. Teacher vs. Student.

4. Discussion and Conclusions

Based on the empirical evidence presented in this study, we characterize the current state of vocational computer curriculum reform in China as a transition from theoretical exploration to practical implementation, albeit with distinct structural imbalances. We observe a "Structural Stratification" in curriculum dependency that highlights a significant disparity in AI integration depth. We find that application-layer courses, such as Python Programming, have achieved a high degree of semantic coupling with AI technologies, whereas infrastructure-layer courses like Operating Systems remain relatively isolated. We interpret this pattern as an indication that the current reform has primarily focused on the adoption of upper-level tools, while the deep integration of AI logic into the foundational computer science curriculum remains a challenge to be addressed in the next phase of development.

Temporally, we identify a notable shift in the focus of educational reform. We view the structural divergence between macro and micro dimensions—particularly the decisive surge of micro-practice attention post-2022—as quantitative evidence that the discourse has moved from macroscopic "Top-level Design" to microscopic "Pedagogical Practice." We note that this trend coincides with the emergence of Generative AI, suggesting that the tangible capabilities of new technologies are driving educators to focus more on specific implementation details—such as code debugging, prompt engineering, and case design—rather than abstract policy discussions. We argue that this shift represents a "Pragmatic Turn" in the field, where the value of reform is increasingly measured by its applicability in actual teaching scenarios.

Finally, through the pronounced structural asymmetry of subject attention, we reveal a critical characteristic of the current reform ecosystem: a high concentration of focus on "Technology" and "Students," contrasted with the marginalization of "Teachers." We indicate that while significant attention is paid to the acquisition of new technologies and the improvement of student learning experiences, the semantic weight assigned to teacher development and digital literacy is comparatively weaker. Since we recognize teachers as the critical mediator between technology and students, we suggest that the sustainability of the reform may depend on closing the attention gap regarding faculty capabilities in future practices. We also acknowledge that this study is bounded by its data source and methodological scope. Consequently, caution should be exercised when generalizing these findings to vocational education systems in other cultural or economic contexts.

Acknowledgements

This work was funded by the Phase III Program (Grant No. 6025310002Q) of the Institute for Economic and Social Development at Shenzhen Polytechnic University.

References

[1] Miao, F., & Holmes, W. (2021). Artificial intelligence and education. Guidance for policy-makers.

[2] Agbo, F. J., Olivia, C., Oguibe, G., Sanusi, I. T., & Sani, G. (2025). Computing education using generative artificial intelligence tools: A systematic literature review. Computers and Education Open, 100266.

[3] Zhang, H. (2025). The Research on the Construction Path of Artificial Intelligence Curriculum System for Higher Vocational Undergraduate Education in the New Era. International Educational Research, 8(2), pp. 99-107.

[4] Prayuda, M. S., Ginting, F. Y. A., & Tamba, L. (2026). Tracing Two Decades of Artificial Intelligence in Education: A Bibliometric Analysis of Trends, Themes, and Future Directions (2000–2025). European Journal of Educational Research, 15(1), 285-304.

[5] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

[6] Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(6), e1332.