Proceedings of CONF-FMCE 2025 Symposium: AI and Machine Learning Applications in Infrastructure Engineering
DOI: 10.54254/2755-2721/2025.MH25386

Technology-Driven Development and Application of Engineering Cost Databases

Shaoyang Zhang

Department of Management, Tianjin University of Technology, Tianjin, China
elmwotui@krae.edu.kg

Abstract. This paper systematically explores the development and application of engineering cost databases driven by technology. Confronting persistent challenges in traditional cost management, such as data fragmentation, over-reliance on standardized quotas, and inefficiency. The integration of technologies: big data, artificial intelligence (AI), cloud computing, blockchain, the Internet of Things (IoT), and Building Information Modeling (BIM) provides robust momentum for database construction. Specifically, it elaborates on key technological pathways for building such databases, encompassing: the integration and standardization of multi-source heterogeneous data; the utilization of data lakes, Natural Language Processing (NLP), knowledge graphs, and intelligent analytical prediction models based on algorithms like Random Forest and Long Short-Term Memory (LSTM); dynamic update mechanisms; and integration with cloud platforms and real-time data acquisition. Practical case studies, including Shandong Province's initiatives, the "Smart Cost" system in Chancheng District, Guangzhou, the material price database of Rizhao Municipal Housing and Urban-Rural Development Bureau, and Yema Technology's AI database, demonstrate the efficacy of these technological applications. Finally, the paper anticipates the future trajectory of engineering cost databases towards increasing intelligence, emphasizing their pivotal role in supporting intelligent construction within international projects and achieving objectives related to engineering intelligence and cost datafication. It posits that such databases will become the core engine driving the construction industry's transformation towards greater refinement, intelligence, and high-quality development.

Keywords: Engineering cost database, Technology-driven, Dynamic cost control, Data integration and standardization, Construction industry digital transformation.

1. Introduction

The rigidity of traditional management models has led domestic cost consulting institutions to excessively rely on standardized quota pricing and government directives, resulting in a lack of independence and professionalism. This rigidity also triggers a vicious cycle where institutions struggle to cultivate innovative cost management methodologies, impeding the integration of advanced technologies like Building Information Modeling (BIM) and blockchain. Moreover, the absence of dynamic data updating mechanisms exacerbates information lag, rendering traditional quota systems increasingly irrelevant in the era of digital construction. Engineering cost databases,

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.MH25386

by contrast, leverage machine learning algorithms to establish predictive models, enabling proactive risk assessment for cost overruns and fostering data-driven collaboration among stakeholders across the project lifecycle. This causes prices to diverge from actual market conditions, which not only suppresses corporate competitive vitality but also hinders their ability to compete internationally. An engineering cost database, a system for storing and managing project cost-related data, has extensive applications and is crucial for engineering management. Such databases serve not only as cost management tools but also as core infrastructure for enhancing industry transparency and facilitating scientific decision-making. With the integration of big data and AI technologies, their application is progressively shifting from "post-facto recording" to "real-time prediction." This enhances the efficiency of cost estimation and budget preparation, optimizes bidding support, enables more rapid and accurate "real-time monitoring" and "change management" to improve dynamic cost control, and allows for stable market trend analysis and forecasting. Through continuous digital upgrades and AI applications in cost management data, they are becoming a critical pillar of the engineering industry's digital transformation [1]. However, engineering cost data suffers from severe fragmentation and inconsistent, dispersed sources, leading to challenges such as difficulties in integrating multi-source heterogeneous data. Consequently, the operational costs associated with utilizing these databases remain high.

2. Development background and core value of engineering cost database

The development of engineering cost databases originates from the pain point that the traditional quota model lags behind market dynamics. With the expansion of engineering construction scale, increased complexity, and the deepening of market-oriented reforms, the industry urgently needs more efficient and accurate cost determination and control methods. The development of information technology provides a technical basis for the collection, storage, analysis, and sharing of massive engineering data such as cost indicators. Meanwhile, the requirements for refined and dynamic cost management, as well as the needs of governments and enterprises to improve investment efficiency, jointly promote the establishment and application of professional engineering cost databases, aiming to achieve rapid updates of resource prices, intelligent analysis of historical data, and provide support for scientific decision-making [2]. Take Chancheng District, Foshan City as an example, its "Smart Cost" system integrates multi-stage data such as estimation, budget, and settlement through AI and big data technologies, realizing full-process digital management and improving audit efficiency by more than 30%. The Engineering Cost Reform Work Planissued by the Ministry of Housing and Urban-Rural Development clearly requires canceling quota price limits, promoting market price inquiry and data accumulation, further highlighting the policy necessity of database construction [3].

3. Key technologies and methods for database construction

3.1. Innovative opportunities driven by technology

The integration of big data, cloud computing, and AI technologies injects new momentum into database construction. For example, the AI database developed by Yema Technology uses natural language processing (NLP) technology to standardize unstructured data and realizes automatic collection and prediction of cost indicators through multi-model training, significantly enhancing data application value. The cloud platform built by the Guangzhou Construction Engineering Cost

DOI: 10.54254/2755-2721/2025.MH25386

Management Station integrates third-party real-time data, releases high-precision information prices, and promotes the construction of industry supervision and supplier credit systems [4].

3.2. Data integration and standardization

The integration of multi-source heterogeneous data relies on a unified framework. The data exchange model based on Industry Foundation Classes (IFC) standards can solve the data fragmentation problem in design, construction, and operation links. For example, the Guangzhou Cost Management Station integrates multi-channel data such as government projects and market transactions through a cloud platform to achieve dynamic updates of material prices. Data cleaning and knowledge graph technologies are used to process unstructured texts and drawings, improving data usability [5].

Driven by data integration and standardization technologies, the availability of engineering cost databases is gradually improving. Distributed storage, parallel computing, and data lake technologies not only support the storage and processing of massive heterogeneous data but also break the capacity limitations of traditional databases. In terms of artificial intelligence and machine learning, natural language processing, computer vision, and automated data cleaning algorithms can automatically extract unstructured data, identify outliers in data, and achieve intelligent classification and labeling.

With the continuous improvement of big data algorithms, driving technologies such as cloud-native technologies (Amazon Web Services, Azure), microservices, and APIs gradually enable cross-platform real-time data synchronization (e.g., Enterprise Resource Planning, BIM systems, supplier quotation platforms), supporting dynamic data updates and global access. In blockchain technology, smart contracts and distributed ledgers ensure the traceability of data sources, prevent tampering, and enhance trust in multi-party collaboration. BIM, IFC standards, and the Omni Class classification system have gradually become industry-standard tools, which can uniformly define bill of quantities, cost codes, and project phase definitions to achieve cross-project data compatibility. In the era of information development and progress, the driving technologies provided by the Internet of Things and real-time data collection (such as sensor networks, 5G communication, edge computing) can collect construction site data in real time and dynamically update cost models.

3.3. Intelligent analysis and dynamic updating

The prediction models of intelligent analysis engineering cost databases are divided into random forest algorithms and LSTM (Long Short-Term Memory) neural networks, which take data as the core and can accurately capture market fluctuation trends, better demonstrating the planning of dynamic cost models in data analysis using random forest algorithms and LSTM neural networks. In random forest algorithms and LSTM neural networks, dynamic data updating is particularly important. In data integration and processing technologies, structured and unstructured data processing, and multi-level data collection can optimize data storage and retrieval, forming data flows from projects to enterprises and then to the industry, enhancing comparability. Prediction machine learning models can narrow the error range, and time series analysis can combine market fluctuation data to predict material price trends and guide the optimization of procurement strategies [2]. Use integrated modeling technologies to embed BIM and 5D models, while performing 3D simulation and optimization of data, conducting semantic analysis and intelligent retrieval, constructing relevant data knowledge graphs, and accurately carrying out dynamic monitoring and optimization models. Cloud computing for dynamic updating supports global data centralized

DOI: 10.54254/2755-2721/2025.MH25386

processing. For example, the "Smart Cost" system in Chancheng District automatically collects cost items through AI, enabling dynamic data iteration and real-time data updates by uploading real-time data to the "Smart Cost" system [5].

3.4. Core advantages and applications at the technical level

The construction of the database centrally stores various engineering cost-related data, realizing data accumulation and sharing, including project costs, material prices, labor costs, etc., facilitating data sharing among different departments and project team members, avoiding repetitive labor, and improving work efficiency [6]. Among them, dynamic model planning provides cost control and prediction. By analyzing historical data, it can provide accurate cost estimation and budget control basis for new projects, help predict possible cost overrun links, and take preventive measures. The algorithm analysis and prediction of the database can improve the accuracy of pricing. The rich engineering cases and price information in the database can make pricing closer to reality, reduce errors, and improve the accuracy and rationality of engineering costs. At the same time, it supports decision-making by providing data support for project decisions. Managers can make more informed investment, bidding, contract signing, and other decisions based on historical data and analysis results.

In the database application stage, the investment estimation stage can use the investment data of similar projects in the database, adjust them combined with the characteristics of new projects, and quickly and accurately estimate the investment scale.

In the design stage, designers can refer to the cost indicators in the database to optimize the design scheme and control the cost in the design stage. For example, by comparing the costs of different structural forms, an economically reasonable design is selected.

In the bidding stage, the bidder can reasonably prepare the bid control price based on the market price information and the winning prices of previous projects in the database. Bidders can also use the database for bid price analysis and formulate competitive bidding strategies.

In the construction stage, real-time monitoring of project costs is carried out, comparing the actual costs with the standard costs in the database, timely identifying cost deviations, and taking corrective measures. At the same time, in terms of material procurement, engineering changes, etc., price negotiations and cost control are carried out with the help of database information.

In the completion settlement stage, refer to the settlement data of similar projects in the database to review the rationality of the completion settlement, prevent overestimation and overcalculation, and ensure the truthfulness and accuracy of the settlement results.

4. Challenges and countermeasures

4.1. Existing problems

Although the integration of big data, cloud computing, and AI technologies has brought new vitality to the database, the resulting uneven data quality leads to insufficient integrity due to diverse sources, requiring the establishment of cleaning and verification mechanisms to make up for existing deficiencies. The integration of multi-source heterogeneous data in the database also brings lagging data standardization. The lagging data standardization and regional quota differences hinder data reuse, and engineering cost databases need to unify coding rules (such as GB/T 51129-2015). With the continuous development of informatization, the security risks faced by engineering cost

DOI: 10.54254/2755-2721/2025.MH25386

databases need to be addressed urgently. The cloud storage of data is vulnerable to attacks, requiring strengthened encryption technologies and permission management [7].

4.2. Solution paths

In the era of data informatization, more advanced data technology optimization can be realized. In the process of data optimization, blockchain technology can be used to ensure data traceability. For example, Yema Technology has directly improved the processing efficiency of non-standard data through NLP data optimization [8]. Policy promotion has driven enterprise engineering data, providing reform pilots for relevant enterprises. For example, Shandong Province proposed the "Data Elements x" strategy to support enterprises in participating in market-oriented reform pilots.

In terms of database security, at the technical protection level, deploy multi-layer firewalls, intrusion detection systems, and data encryption technologies, and implement strict access control and identity authentication mechanisms; at the management specification level, establish a hierarchical authorization system, regularly conduct security audits and vulnerability scans, and formulate data backup and disaster recovery plans; at the dynamic monitoring level, real-time monitor abnormal operation behaviors, track risk sources through log analysis, and promptly block attack paths. At the same time, compliance assurance is also an important reason, including abiding by data security regulations, strengthening personnel safety awareness training, and implementing main body responsibilities.

In today's rapidly developing technology, ecological co-construction provides a shared platform for engineering cost databases. Industry-level shared platforms are constructed to realize industrial data exchange and communication. For example, Cost-communication covers 326 cities nationwide, promoting data circulation and value mining [9].

5. Conclusion

Through the research on the construction and application of engineering cost databases, the following core conclusions are drawn: This technology effectively solves the long-standing problem of fragmented data in the engineering construction field, replacing the traditional inefficient management model relying on personal experience and scattered documents. Its standardized and structured data storage method constructs a unified and authoritative data foundation, eliminates information barriers, and provides a reliable basis for scientific decision-making; it significantly improves work efficiency and result quality, realizes the automatic accumulation and intelligent reuse of historical data, greatly shortens the time for repetitive price inquiries and calculations, and through unified data calibers and a perfect traceability mechanism, significantly enhances the accuracy, comparability, and audit reliability of cost results, effectively reducing dispute risks; it lays a key foundation for the intelligent upgrading of the industry.

Looking to the future, as an indispensable core data support, the engineering cost database will strongly promote the implementation of cutting-edge technologies such as in-depth BIM application, AI intelligent cost review, and automated quantity calculation, and promote the transformation of cost work from labor-intensive to knowledge-intensive and technology-intensive, continuously leading the digital transformation process of the engineering construction field.

In summary, the engineering cost database has evolved from an auxiliary tool to a strategic infrastructure driving the digital transformation and high-quality development of the engineering construction field. Its future development will deeply integrate cutting-edge technologies such as AI, big data, and blockchain, playing a key role as a "data hub" in improving the precision of micro-

DOI: 10.54254/2755-2721/2025.MH25386

project management, as well as in macro resource allocation, industry governance model innovation, and smart city construction. Sustained investment in the research, development, and application of this field is not only an inevitable choice for enterprises to reduce costs and increase efficiency but also the core path for the country to improve engineering construction quality, optimize investment benefits, and achieve sustainable development of the construction industry.

References

[1] Research on the Development of Project Cost Informatization in the Era of Big Data. Journal | Journal on Big Data. Volume 2, Issue 2, 2020. PP 63-70.
[2] Rizhao Housing and Urban-Rural Development Bureau. Engineering Cost Data of State-owned Funds Projects Helps Deepen Market-Oriented Reform of Engineering Costs. 2024.
[3] Research on the Cost Prediction Model of Construction Projects Based on the Support Vector Regression Machine. Kong Xiangpeng. BASIC & CLINICAL PHARMACOLOGY & TOXICOLOGY. 2020.
[4] Liu Datong, Zhao Bin. Research on the Construction of Cost Databases under the Background of Engineering Cost Reform. Engineering Construction Standardization, 2023.
[5] Financial Bureau of Chancheng District, Foshan City. Application Report of Smart Cost System. 2024.
[6] Dai Duanming. Construction of Engineering Cost Database Based on Big Data Technology. 2021: 79-83.
[7] Tang Anbin. Research and Application of Cost Index Database for Engineering Cost Consulting Enterprises. Southwest University of Science and Technology, 2022.
[8] Yema Technology. AI Database Technology White Paper. 2024.
[9] China Construction Engineering Cost Management Association. Feasibility Study on Green and Low-Carbon Building Cost Database. 2024.