Knowledge Graph Reasoning for Risk Assessment in Green Industrial Project Finance

Authors

Yuki Nakamura[1], Haruka Sato[2]*

Affiliations

Graduate School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan

***Corresponding author:** h.sato@kyoto-u.ac.jp

Abstract

Green industrial project finance supports renewable energy facilities, low-carbon manufacturing upgrades, pollution-control equipment, energy-saving renovation, and circular economy projects. These projects involve long investment cycles, policy subsidies, technical uncertainty, environmental compliance pressure, and complex stakeholder relationships. Traditional project risk assessment often focuses on financial ratios and collateral value, but it does not sufficiently capture policy dependency, environmental performance, supplier reliability, and sponsor credit risk. This study develops a knowledge graph reasoning model for risk assessment in green industrial project finance. The proposed method builds a project-centered knowledge graph containing project sponsors, engineering contractors, equipment suppliers, subsidy policies, environmental permits, carbon-reduction indicators, loan contracts, repayment schedules, and related enterprise networks. A graph neural encoder learns project risk representations, while a rule-based inference module identifies risk chains from delayed subsidies, weak environmental compliance, contractor disputes, technology performance gaps, and sponsor financial stress. The evaluation dataset contains 8,760 green industrial projects, 24,300 sponsor enterprises, 6,420 contractors, 18,900 equipment suppliers, 31,600 policy-subsidy records, 52,000 financing contracts, and 4,180 project-risk events over 50 months. The proposed model reduces median early-warning time from 96 days to 41 days before project repayment deterioration. It identifies 3,260 policy-dependent risk paths and 2,140 contractor-related delay chains. Portfolio simulation shows that prioritizing projects by graph-inferred risk exposure reduces expected overdue financing amount by 132 million RMB during the validation period. Full quarterly assessment is completed in 9.8 minutes. The findings indicate that knowledge graph reasoning can provide more interpretable and relationship-aware risk assessment for green industrial project finance.

Keywords: Green industrial finance; project risk assessment; knowledge graph reasoning; graph neural network; environmental compliance; subsidy risk; industrial financial risk.

1. Introduction

Green industrial project finance plays an important role in supporting renewable energy development, low-carbon industrial upgrading, pollution-control facilities, energy-efficiency improvements, and circular economy projects. These projects are often characterized by large initial investments, long payback periods, strong policy dependence, strict environmental requirements, and complex contractual relationships among sponsors, contractors, suppliers, financial institutions, and public authorities. As a result, their risk profiles are influenced not only by project cash flows and sponsor solvency but also by policy changes, subsidy implementation, environmental compliance, technology performance, construction progress, supplier reliability, and stakeholder coordination [1,2]. Traditional project finance assessment mainly relies on financial ratios, projected cash flows, debt-service capacity, collateral conditions, and sponsor credit information [3]. These indicators are essential for evaluating repayment capacity, but they often provide only a partial view of project risk because many non-financial events and cross-entity relationships are evaluated separately. Recent research on cloud infrastructure security has also emphasized the value of open-source frameworks that integrate multiple monitoring components and security mechanisms within a unified architecture [4]. This framework-based perspective provides a useful reference for green project risk management, where risk information is distributed across financial, operational, environmental, policy, and contractual systems and therefore requires integrated rather than isolated analysis [5,6].

Recent studies have increasingly incorporated relational, transactional, and operational data into financial risk assessment. Information such as subsidy records, invoices, payment transactions, ownership relationships, supplier links, environmental permits, and stakeholder interactions can provide early signals of project deterioration that are not directly visible in conventional financial statements [7,8]. Machine learning and network-based methods have improved the detection of nonlinear relationships and hidden dependencies among multiple risk factors. However, most existing approaches still analyze financial indicators and relational structures through separate modeling processes or use relatively limited data sources. Such designs are not well suited to green industrial projects because project risk often develops through a sequence of connected events [9,10]. A delayed subsidy payment may weaken project liquidity, which may then delay contractor payments, disrupt equipment delivery, postpone construction milestones, and eventually increase debt-servicing pressure [11,12]. Similarly, weak environmental compliance may lead to permit restrictions, additional remediation costs, operational delays, or financing constraints [13,14]. Existing models provide limited support for representing these linked events as a continuous risk process and often offer insufficient explanations of how risk originates and propagates across project participants [15,16]. Knowledge graphs provide a suitable structure for integrating heterogeneous information from green industrial projects because they represent entities, attributes, transactions, and events through explicit relationships [17,18]. Sponsors, contractors, suppliers, financing institutions, subsidy programs, environmental permits, carbon indicators, equipment contracts, loan agreements, repayment schedules, and project milestones can be organized within a unified project-centered graph. This representation allows project risk to be examined from both entity-level conditions and inter-entity dependencies [19,20]. Graph-based learning can further capture latent relationships among connected participants and identify structural patterns that are difficult to observe through tabular variables alone [21,22]. When combined with rule-based reasoning, the model can trace risk chains associated with delayed subsidy disbursement, contractor disputes, supplier disruption, environmental non-compliance, technology underperformance, abnormal cost increases, and sponsor financial stress. Such reasoning is particularly useful for green industrial finance because a single adverse event may affect several parties and gradually spread from project execution to financing and repayment performance [23,24].

This study develops a knowledge graph reasoning approach for risk assessment in green industrial projects. The proposed framework constructs a project-centered knowledge graph that integrates sponsors, contractors, suppliers, subsidy policies, environmental permits, carbon-related indicators, financing arrangements, project milestones, loans, and repayment schedules. Graph-based representations are used to capture structural dependencies among project participants, while rule-based reasoning is applied to identify interpretable risk pathways generated by policy delays, compliance problems, contractor disputes, technology gaps, supplier instability, and sponsor financial pressure. The empirical analysis uses a dataset covering 8,760 green industrial projects, 24,300 sponsors, 6,420 contractors, 18,900 suppliers, and 4,180 confirmed risk events over a 50-month period. The study aims to improve the timeliness and interpretability of project risk identification by moving beyond conventional financial-ratio assessment toward multi-source and relationship-based analysis. It also seeks to reveal how individual risk events develop into broader project-level problems and to provide traceable evidence for early-warning decisions. From a practical perspective, the proposed approach can help financial institutions prioritize high-risk projects, strengthen portfolio supervision, and improve the allocation of monitoring resources. It can also support project managers in identifying vulnerable contractual and operational links and provide regulators with a more systematic view of policy, environmental, and financial interactions. In this way, the study extends green project finance risk assessment from static financial evaluation to dynamic relational reasoning and provides a practical framework for improving the transparency, efficiency, and explainability of risk management in green industrial investment.

2. Materials and Methods

2.1 Samples and Study Area

This study used a dataset of 8,760 green industrial projects over 50 months. The sample included 24,300 project sponsors, 6,420 contractors, 18,900 equipment suppliers, 31,600 policy-subsidy records, 52,000 financing contracts, and 4,180 confirmed project-risk events. Projects with missing sponsor identifiers, incomplete financing records, unclear approval dates, or incomplete status information were removed. The projects cover renewable energy, low-carbon manufacturing, pollution-control equipment, energy efficiency upgrades, and circular economy initiatives. These projects were chosen because green industrial investments involve long cycles, policy dependence, environmental compliance, and close links between sponsors, contractors, suppliers, and financiers.

2.2 Experimental Design and Control Settings

The experiment tested whether the proposed knowledge graph reasoning method improves project risk assessment compared with conventional approaches. The proposed method was used as the experimental group. It combined sponsors, contractors, suppliers, subsidy policies, environmental permits, carbon indicators, loan contracts, repayment schedules, and related enterprise networks into a project-centered knowledge graph. Two control groups were used for comparison. The first relied on financial-ratio scoring using leverage, collateral, repayment schedule, sponsor profitability, and project cash-flow coverage. The second used a tree-based classifier with the same project and sponsor-level variables. These control groups were chosen because financial scoring and tree-based classification are standard in project loan review. All methods used the same training and testing periods.

2.3 Measurement Methods and Quality Control

Project risk was measured using repayment deterioration, delayed subsidies, environmental compliance issues, contractor disputes, equipment delays, technology performance gaps, and sponsor financial stress. Subsidy records measured policy dependency and payment delays. Permits and environmental compliance documents confirmed adherence to standards. Contractor and supplier records identified delays or disputes. Loan contracts and repayment schedules quantified financing pressure. Enterprise names, sponsor identifiers, project IDs, policy codes, and contract numbers were standardized. Duplicate or inconsistent records were removed. Extreme values in financial variables were checked using percentiles and business rules. A time-based split was applied to avoid using future risk events in training.

2.4 Data Processing and Model Formulation

Project data were organized by quarter and project type. Continuous variables were standardized within project categories to reduce variation due to project size or industry. Missing continuous values were replaced with the median for the same category and quarter. Missing categorical values were assigned to a separate group. The project knowledge graph at period t is defined as:

$$G_t=(V_t,E_t,X_t,R_t),$$

where V_t represents projects, sponsors, contractors, suppliers, policies, permits, and loan nodes; E_t represents links between these nodes; X_t denotes node attributes; and R_t denotes link types, including sponsorship, contracting, supply, subsidy, permit, loan, and repayment links. The project risk score was calculated as:

$$P_i=\sigma(Wh_i+b),$$

where P_i is the estimated risk of project i, h_i is the project representation from attributes and linked nodes, W and b are estimated parameters, and $\sigma(\cdot)$ is the sigmoid function. Early-warning lead time was defined as:

$$L_i=T_i^{risk}-T_i^{warning},$$

where T_i^{risk} is the date of confirmed project risk and $T_i^{warning}$ is the first date the risk score exceeded the threshold. A higher L_i indicates earlier detection.

2.5 Evaluation and Robustness Analysis

The method was evaluated based on early-warning time, identification accuracy, risk-chain clarity, and portfolio-level impact. Early-warning time was measured by the median number of days between first warning and confirmed repayment deterioration. Accuracy metrics included AUC, precision, recall, F1-score, and top-risk capture rate. Risk-chain clarity was measured by the number of identified paths related to delayed subsidies, environmental compliance, contractor disputes, equipment delays, technology gaps, and sponsor financial stress. Portfolio-level impact was evaluated by the reduction in expected overdue financing when projects were ranked by estimated risk. Robustness tests included project type, investment size, policy-dependency level, and warning threshold. Additional tests removed subsidy records, permits, contractor links, supplier links, and rule-based reasoning individually to measure their contributions.

3. Results and Discussion

3.1 Early-Warning Effectiveness

The proposed knowledge graph reasoning method provided earlier warnings of project repayment issues than the financial-ratio baseline. The median early-warning time decreased from 96 days to 41 days, showing that risk signals can be detected well before confirmed distress events. Conventional methods rely mainly on leverage, collateral, and cash-flow ratios, which often identify problems only after financial stress appears. By combining sponsors, contractors, suppliers, subsidy policies, environmental permits, carbon indicators, loans, and repayment schedules, the method captures multiple risk sources that are often overlooked. These findings agree with previous research showing that integrating relational information improves project risk assessment. Fig. 1 illustrates a multi-view graph used to link different financial connections and support more complete risk evaluation [25,26].

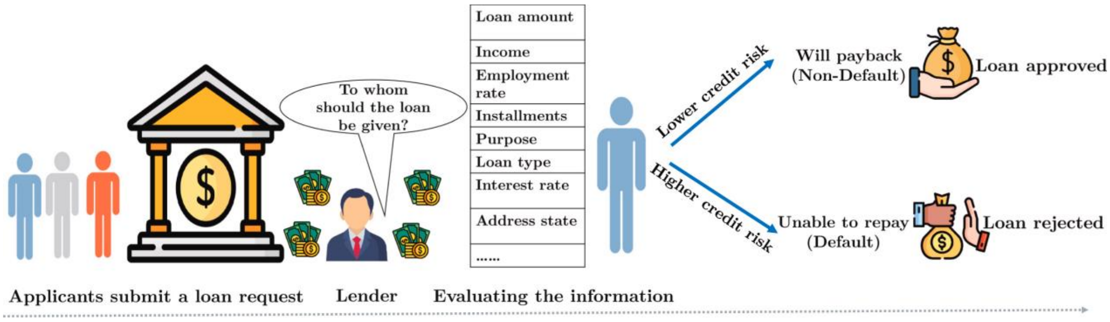

Fig. 1. Project graph linking sponsors, contractors, suppliers, and policy records for risk assessment.

3.2 Policy and Compliance Risk Paths

The method identified thousands of risk paths related to delayed subsidies, weak environmental compliance, contractor disputes, and sponsor financial stress. Policy delays were common in projects with high subsidy dependence, increasing financing pressure and delaying construction milestones. Weak environmental compliance or permit issues often preceded repayment deterioration. Earlier studies focused on single connections or simple network effects and did not integrate policy and compliance factors. The risk paths in this study demonstrate how policy and compliance failures propagate through sponsor and contractor links, providing a clearer understanding of project risk dynamics beyond traditional financial measures [27,28].

3.3 Contractor and Supplier Risk Contributions

Contractor delays and supplier issues were major contributors to project risk. Projects with linked contractor disputes or delayed equipment deliveries had higher risk scores and earlier warnings than projects without such links. This shows that operational factors such as construction delays and supplier instability are critical for assessing project health. Fig. 2 shows a network perspective where connected records reveal complex risk propagation in financial data. Embedding stakeholder and transaction connections in risk assessment captures interdependencies that conventional metrics may miss [29,30].

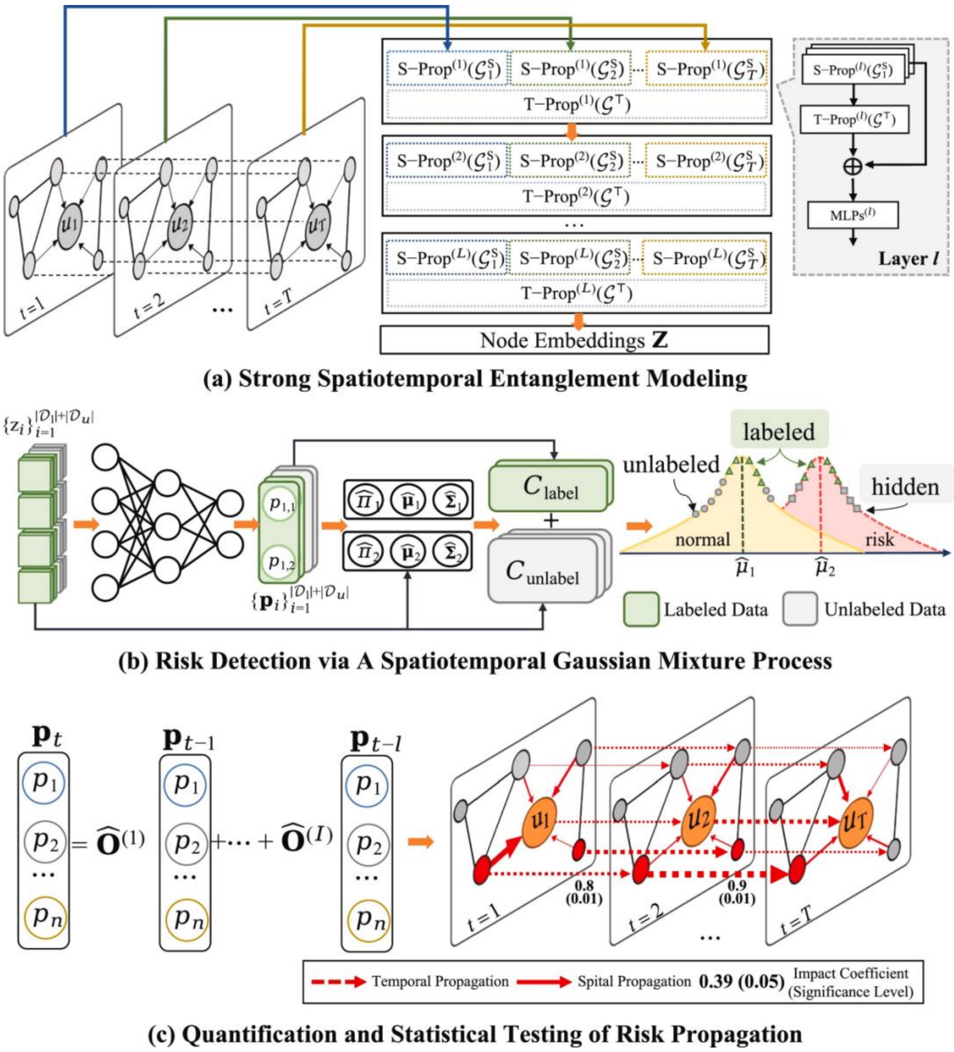

Fig. 2. Risk paths showing projects affected by subsidy delays, contractor disputes, and supplier problems.

3.4 Comparison with Previous Studies and Robustness

Compared with earlier research, this method advances project risk assessment by integrating policy, compliance, sponsor, contractor, and supplier information into a project-centered structure. Prior approaches often relied only on firm-level financial indicators or limited network connections. Robustness tests across project categories, investment sizes, and policy dependence levels showed consistent improvements. Additional tests removing subsidy records, permit links, or contractor connections confirmed that each data source contributes to early detection and risk path clarity. These results indicate that knowledge graph reasoning can provide earlier warnings, more interpretable risk pathways, and more effective prioritization of green industrial projects than conventional financial methods.

4. Conclusion

This study proposed a knowledge graph reasoning method for assessing risk in green industrial project finance. The method combined project sponsors, contractors, suppliers, subsidy policies, environmental permits, carbon indicators, loan contracts, and repayment records into a project-centered graph. Results showed that the method provided earlier warnings than conventional financial-ratio approaches and identified key risk paths related to subsidy delays, weak compliance, contractor disputes, supplier problems, and sponsor financial stress. The main contribution is that it moves risk assessment beyond single financial indicators to include policy, environmental, and stakeholder connections. This provides a clearer understanding of how project risk arises and spreads. The method can help banks, project managers, and regulators in early warning, project screening, portfolio review, and supervision of green finance. However, the study has limitations. Results depend on the completeness and accuracy of policy, permit, contract, and repayment records. Some informal disputes or technical issues may not be fully captured. Future studies should test the method in more regions and project types and incorporate more real-time construction and environmental data.

References

1. Yin, J., Rao, H., & Huang, Y. (2026, March). Dynamic Modeling and Heterogeneity Analysis of Platform User Behavior Time Series. In 2026 International Conference on AI in Education Technology and Applications (AIETA) (pp. 30-33). IEEE.

2. Kirui, N. K., Rambo, C. M., & Luketero, S. W. (2025). Moderating effects of project environment on financial risk management and safety of road construction projects in Kenya. International Journal of Project Organisation and Management, 17(3), 327-352.

3. Zhang, Z. (2026). A Study on the Identification of Manipulative Design in Subscription and Payment Interfaces of Digital Consumer Platforms and Its Behavioral Effects. Available at SSRN 6734760.

4. Shao, W. (2026). Design and Implementation of an Open-Source Security Framework for Cloud Infrastructure. arXiv preprint arXiv:2604.03331.

5. Qi, C., & Qiao, X. (2026). Building Reliable AI Systems: A Framework That Bridges Architecture and Operations. Available at SSRN 6795298.

6. Al-Qudah, S., Bargues, J. L. F., & Gisbert, P. S. F. (2026). A Methodological Proposal for Risk Management in Construction Projects in Small and Medium-Sized Enterprises in Emerging Economies: The Case of Jordan.

7. Xu, T., Zhu, W., & Zhang, J. (2026). Stability and Consistency of Explainable Deep Learning Methods in Credit Risk Assessment. Available at SSRN 6893938.

8. Yin, L., Tong, Y., Xie, R., Zhang, Z., Islam, Z. H., Zhang, K., ... & Wang, B. (2024). Targeted NAD+ repletion via biomimetic nanoparticle enables simultaneous management of intimal hyperplasia and accelerated re-endothelialization: a proof-of-concept study toward next-generation of endothelium-protective, anti-restenotic therapy. Journal of Controlled Release, 376, 806-815.

9. Xiong, W., Zeng, Y., & Guo, Y. (2026). A Study on the Characterization of Building MEP System States and the Learning of Behavioral Patterns Driven by Large-Scale Operational Data. Available at SSRN 7121883.
10. Rahman, M. S. (2025). Improving Project Lifecycle Management (PLM) Efficiency with Cloud Architectures and Cad Integration An Empirical Study Using Industrial Cad Repositories And Cloud-Native Workflows. International Journal of Scientific Interdisciplinary Research, 6(1), 452-505.
11. Hong, Z., Xu, T., & Chen, H. (2026). Automated Detection and Reclamation of Orphaned Compute and Storage Resources in Cloud Infrastructure. Available at SSRN 7116258.
12. Wu, J., Wu, D., Zhang, J., & Peng, Y. (2026). Generative AI Feedback in Junior High School Artistic Creation Evaluation. Available at SSRN 7244838.
13. Huang, J., Xu, T., Yang, J., & Yin, J. (2026). A Graph Neural Network Approach for Financial Data Validation and Risk Identification in Large-Scale Intercompany Transactions. Available at SSRN 7175058.
14. Kussaiyn, M., & Potluri, R. M. (2026). Integrating ESG factors into cost forecasting for sustainable project management: Empirical evidence from Kazakhstan. World Development Sustainability, 100279.
15. Huang, J., Xu, T., Yang, J., & Yin, J. (2026). Blockage Early Warning and Financial Impact Quantification in the Monthly Closing Process of Fintech Companies. Available at SSRN 7179198.
16. Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. Advances in Neural Information Processing Systems, 35, 12800-12812.
17. Gu, X. (2026). Identifying Causal Effects and Analyzing Heterogeneity of User Growth Interventions on Digital Platforms: Evidence from Large-Scale Behavioral Data. Available at SSRN 6809181.
18. Peeris, M. P., Baryannis, G., & Papadakis, E. (2026). Advancing Sustainable Supply Chains Through Knowledge Graph Completion and Graph-Based Artificial Intelligence. Sustainability, 18(6), 2825.
19. Zhao, J., Fan, J., & Li, L. (2026). A Study on an Explainable Causal-Enhanced LLM Agent for Predicting the Forming Quality of Automotive Component Materials.
20. You, S. (2026). Verifiable Audit Mechanisms in AI Compliance Automation: Scalability. Available at SSRN 6547458.
21. Wu, Y., & Su, D. (2026). An Evaluation of Parental Question Characteristics and the Quality of AI Recommendations in Family Counseling for Children with Autism.
22. Nguyen, T., Nguyen, K., Nguyen, P., Vo, K. T., Ta, T. T., Nguyen, P., ... & Nguyen-Hoang, T. A. (2025). Graphgru-ed: Leveraging graph-based imputation and deep learning for MOOC completion prediction. IEEE Access, 13, 195272-195294.

23. Jiao, Y., Shi, T., Zhao, B., & Wang, A. (2026). Retrieval-Guided Structured Reasoning and Interpretable Representation Learning for Large-Scale Video–Language Models.
24. Mohamadi, F. (2026). Contracts in Non-recourse Project Finance. In Introduction to Project Finance in Renewable Energy Infrastructure: Sustainable Finance and Investing for a Just Energy Transition to a Low-Carbon Economy (pp. 325-378). Cham: Springer Nature Switzerland.
25. Zhang, G., Xu, Y., Zhang, M., Wang, S., & Lin, N. (2021). The brain network in support of social semantic accumulation. Social cognitive and affective neuroscience, 16(4), 393-405.
26. Bao, Y., & Wang, H. (2026). Edge-to-Cloud Infrastructure Continuum: A Unified Framework for Optimizing Industrial IoT Systems. Available at SSRN 7185578.
27. Alkindi, S. (2026). Decision-making strategies for project management and financial planning: UAE's public sector construction projects (Doctoral dissertation, Anglia Ruskin Research Online (ARRO)).
28. Wang, S., Feng, Y., & Fang, X. (2026, May). A Large Language Model-Enabled Multi-Agent Collaboration Method for Complex Task Solving. In 2026 6th International Symposium on Computer Technology and Information Science (ISCTIS) (pp. 253-256). IEEE
29. Owolabi, O. S. (2025). Network Analysis for Systemic Risk Assessment in Supply Chains: A Cross-Disciplinary Framework Integrating Financial Contagion Models. Journal of Data Analysis and Information Processing, 2025, 13 (3), 23.
30. Bao, Y., Qiu, Y., & Wang, H. (2026). FLARE: A Real-Time Framework for Detecting Malicious Account Activities at Internet Scale. Available at SSRN 7185659.