Environmental Compliance Graph Reasoning for Financing Risk Prediction in Industrial Enterprises

Authors

Anna Meier[1], Dr. Lukas Steiner[2]*

Affiliations

Department of Finance, University of Zurich, Zurich 8006, Switzerland

***Corresponding author:** l.steiner@uzh.ch

Abstract

Environmental compliance has become an important factor in industrial enterprise financing, especially for firms in chemicals, metallurgy, energy, building materials, and heavy manufacturing. Environmental penalties, emission violations, production restrictions, delayed rectification, and pollution-control investment pressure may affect cash flow, credit access, and repayment stability. This study proposes an environmental compliance graph reasoning model for financing risk prediction in industrial enterprises. The model constructs a heterogeneous knowledge graph linking enterprises, environmental penalties, emission permits, production facilities, pollution-control equipment, bank loans, supplier relationships, customer orders, government subsidies, and repayment outcomes. A graph neural network is used to encode enterprise compliance and financing dependency, while a knowledge reasoning module identifies risk chains involving repeated penalties, suspended production, delayed rectification, subsidy withdrawal, supplier disruption, and loan repayment pressure. Experiments are conducted on an industrial compliance-finance dataset containing 57,300 enterprises, 138,000 environmental inspection records, 42,600 penalty events, 96,000 emission-permit records, 710,000 supplier-customer links, 52,400 loan contracts, and 7,860 confirmed financing-risk cases over 50 months. The proposed model shortens median financing-risk warning time from 82 days to 30 days compared with a financial-indicator baseline. Graph reasoning identifies 6,240 compliance-driven risk chains and 2,080 enterprises affected by both environmental penalties and order contraction. The system consolidates 15,700 raw enterprise alerts into 3,890 investigation cases through compliance-path aggregation. Full monthly assessment is completed in 13.9 minutes, with a median inference latency of 47 ms per enterprise node. The results indicate that environmental compliance graph reasoning can improve industrial financing risk prediction by connecting regulatory events, operational constraints, and enterprise credit behavior.

Keywords: Environmental compliance; industrial financing risk; knowledge graph reasoning; graph neural network; emission penalty; credit risk prediction; enterprise risk chain.

1. Introduction

Environmental compliance has become an increasingly important determinant of financing risk for industrial enterprises, particularly in chemicals, metallurgy, energy, building materials, and other pollution-intensive manufacturing sectors. Environmental penalties, emission violations, production restrictions, delayed rectification, and large pollution-control expenditures can directly affect operating costs, production continuity, cash flow, and access to external financing [1,2]. At the same time, environmental supervision is becoming more dependent on digital monitoring, cloud-based data management, and continuous reporting systems [3]. Research on open-source cloud infrastructure has emphasized the importance of secure and reliable digital environments for maintaining the integrity and availability of operational data [4], which is also relevant to the collection and transmission of environmental compliance information. When compliance records, emission data, or regulatory reports are incomplete, delayed, or unreliable, financial institutions may face greater difficulty in assessing the true environmental and operational exposure of a borrower [5,6].

Traditional credit risk assessment methods mainly rely on financial statements, collateral values, historical repayment records, and other borrower-level indicators [7,8]. These variables remain useful for general credit evaluation, but they often fail to capture environmental events that can rapidly change a firm's operating condition. A major environmental penalty may create unexpected cash outflows, while production suspension or capacity restrictions can reduce revenue and weaken debt-servicing ability [9,10]. Delayed rectification may lead to repeated regulatory action, and substantial expenditure on pollution-control equipment may increase short-term liquidity pressure even when it improves long-term compliance [11,12]. Previous studies have identified relationships among regulatory enforcement, corporate environmental performance, and credit outcomes [13,14]. However, many existing approaches focus on individual indicators, static compliance records, or relatively small samples, making it difficult to represent the combined effect of multiple regulatory and operational events on financing risk [15,16]. Graph-based learning and knowledge graph techniques provide a promising way to address these limitations because they can represent enterprises, regulatory events, permits, production facilities, suppliers, customers, contracts, and financial obligations within a unified relational structure [17,18]. Unlike conventional models that treat each enterprise as an isolated observation, graph neural networks can aggregate information from connected entities and identify hidden dependencies through which environmental and financial risks may propagate. For example, an enterprise subject to production restrictions may affect downstream customers and upstream suppliers, while repeated compliance violations can simultaneously influence operating capacity, cash flow, contractual performance, and refinancing conditions. Knowledge graphs can further improve interpretability by linking predicted risk to specific events and relationships rather than producing only an aggregate risk score [19,20]. Despite these advances, environmental compliance information remains insufficiently integrated into industrial financing models. Existing studies rarely combine environmental penalties, emission permits, rectification records, production constraints, pollution-control investments, supplier-customer relationships, and repayment outcomes within the same analytical framework [21,22]. Many predictive models also provide limited explanations of how environmental events evolve into financial stress. This limitation reduces their practical value for banks and other financial institutions, where risk monitoring requires not only an accurate probability estimate but also clear evidence of the events and relationships responsible for deteriorating credit quality [23,24]. The temporal nature of compliance risk is also frequently overlooked. A single penalty may have limited financial impact, whereas repeated violations, prolonged production restrictions, or unresolved rectification requirements may indicate a persistent deterioration in operational stability.

To address these limitations, this study proposes an environmental compliance graph reasoning model for industrial financing risk assessment. The proposed framework constructs a heterogeneous graph linking enterprises, environmental penalties, emission permits, regulatory inspections, rectification records, production restrictions, pollution-control investments, suppliers, customers, financing contracts, and repayment outcomes. Graph-based representation learning is used to capture dependencies among regulatory, operational, and financial entities, while an interpretable reasoning mechanism traces risk pathways through events such as repeated violations, delayed rectification, production suspension, rising environmental expenditure, and weakening business relationships. By incorporating environmental compliance signals into conventional credit assessment, the model aims to identify financing risk before it is fully reflected in financial statements or repayment records. The framework also provides traceable explanations of how specific compliance events contribute to operational disruption and repayment pressure, supporting banks in credit screening, post-loan monitoring, and early intervention. More broadly, the study provides a practical approach for integrating environmental governance information into industrial credit risk management and may help financial institutions better evaluate borrowers operating under increasingly stringent environmental regulation.

2. Materials and Methods

2.1 Sample and Study Area Description

This study analyzed industrial enterprises in sectors with high environmental impact, including chemicals, metallurgy, energy, building materials, and heavy manufacturing. The dataset included 57,300 enterprises observed over 50 months. Firms were selected if they had records on environmental inspections, penalties, emission permits, production facilities, pollution-control equipment, supplier-customer links, bank loans, subsidies, and repayment outcomes. The dataset contained 138,000 inspection records, 42,600 penalties, 96,000 emission permits, 710,000 supplier-customer links, 52,400 loan contracts, and 7,860 confirmed financing-risk cases. The sample covered small, medium, and large enterprises, providing a realistic setting to study how environmental compliance affects financing risk and repayment stability.

2.2 Experimental Design and Control Experiments

The study compared the proposed environmental compliance graph reasoning model with traditional financial-indicator methods. The experimental group used a knowledge graph linking regulatory events, production constraints, supplier-customer networks, and loan outcomes. The control group used a baseline model based on liquidity ratios, leverage, revenue trends, collateral, repayment history, and firm age. Two additional comparisons included a logistic regression model and a tree-based machine learning model using structured enterprise indicators. The data were split chronologically into training, validation, and testing sets. Early-period data trained the models, and later-period data tested whether financing risks could be detected before repayment stress appeared. This design evaluates both predictive performance and early warning capability.

2.3 Measurement Methods and Quality Control

Financing risk was measured using confirmed repayment stress events and related adverse outcomes. A financing-risk case was defined as an enterprise with one or more events, including overdue repayment, credit-limit reduction, loan renewal failure, cash-flow stress, production restriction, delayed rectification, subsidy withdrawal, supplier disruption, or order contraction. Compliance was measured by inspection frequency, penalty count, penalty amount, emission-permit status, rectification delay, and pollution-control equipment status. Production pressure was measured by restricted facilities and their link to active orders. Supplier and customer exposure was calculated as the share of key partners in the enterprise network. Data quality control included entity standardization, removal of duplicates, timestamp checks, and verification of outliers. Missing financial values were replaced with industry-year medians, and missing relational records were kept as sparse links to avoid false connections.

2.4 Data Processing and Model Formulation

All records were converted into a heterogeneous knowledge graph $G=(V,E,R)$, where V is the set of nodes, E the set of edges, and RRR the set of relation types. Nodes represented enterprises, inspection events, penalties, emission permits, production facilities, pollution-control equipment, loans, suppliers, customers, subsidies, and repayment outcomes. Edges represented verified relationships, including enterprise-penalty links, permit ownership, facility operation, equipment installation, loan issuance, supplier-customer transactions, subsidy receipt, and repayment status. Numerical variables were normalized by industry and month, and categorical variables were encoded as embedding vectors. The compliance pressure score for enterprise i was calculated as:

$$C_i=\frac{P_i+\lambda_1 R_i+\lambda_2 S_i}{1+\log(1+A_i)}$$

where P_i is the number of penalties, R_i the number of rectification-delay events, S_i production restrictions, A_i firm assets, and λ_1,λ_2 are weighting coefficients. Node representations were updated using relation-aware aggregation:

$$h_i^{(l+1)} = \sigma\left(W_0^{(l)} h_i^{(l)} + \sum_{r \in R} \sum_{j \in N_i^r} \frac{1}{c_{i,r}} W_r^{(l)} h_j^{(l)} \right)$$

where $h_i^{(l)}$ is the node representation at layer l, N_i^r is the set of neighbors connected by relation rrr, $W_0^{(l)}$ and $W_r^{(l)}$ are trainable weights, $c_{i,r}$ is a normalization factor, and $\sigma(\cdot)$ is the activation function. The final embeddings were used to estimate financing-risk probability and generate interpretable risk chains.

2.5 Model Evaluation and Implementation

Model evaluation focused on prediction accuracy, early warning time, and interpretability. Accuracy was measured using AUC, precision, recall, F1-score, and calibration error. Early warning time was the median number of days between model warning and confirmed repayment stress. Interpretability was assessed by the number and structure of risk chains, including repeated penalties, suspended production, delayed rectification, subsidy withdrawal, supplier disruption, order contraction, and repayment pressure. Alert aggregation was evaluated by comparing raw enterprise alerts with grouped investigation cases. The model was implemented using a graph learning framework with mini-batch training to handle large-scale enterprise networks. Monthly assessment of all enterprises was completed in 13.9 minutes, with a median inference latency of 47 ms per enterprise node, showing the method's efficiency for industrial financing risk monitoring.

3. Results and Discussion

3.1 Financing-Risk Prediction Performance

The environmental compliance graph reasoning model performed better than the financial-indicator baseline, logistic regression model, and tree-based model. This improvement was mainly due to the joint use of compliance records and enterprise links. The financial baseline used liquidity, leverage, revenue change, collateral, and repayment history. These indicators are useful, but they do not fully reflect risks caused by repeated penalties, late rectification, production limits, supplier disruption, and order decline. The proposed model linked regulatory events with facility status, supply-chain relations, loan contracts, and repayment outcomes. As shown in Fig.1, the graph-based model gave more stable results across monthly tests, especially when some financial records were incomplete. This result is consistent with relational credit-risk studies, which show that firm risk may be affected by connected firms and public risk events in enterprise networks [25,26]. The finding suggests that environmental compliance should be treated as a financing-risk factor linked to enterprise operations, rather than as a separate regulatory record.

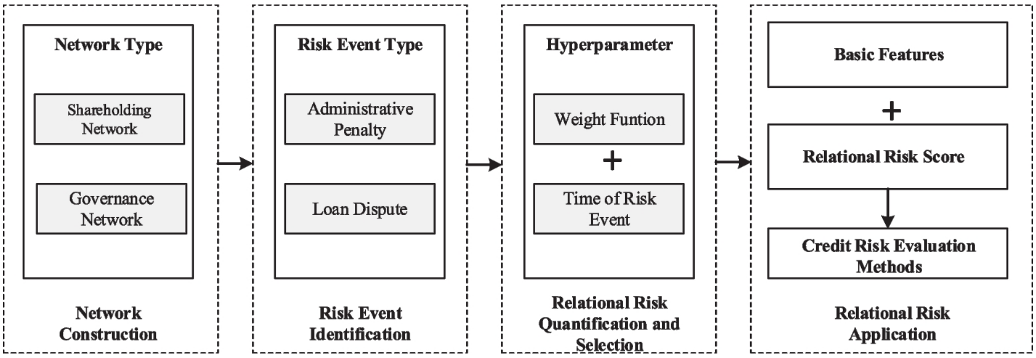

Fig. 1. Financing-risk prediction performance of the environmental compliance graph reasoning model and baseline models.

3.2 Early Warning Effect of Compliance Information

The proposed model reduced the median financing-risk warning time from 82 days to 30 days. This result shows that compliance records can provide earlier risk signals than financial ratios alone. Some firms still had acceptable short-term financial performance, but their inspection records showed repeated violations, late rectification, or production limits. These events later caused weaker order delivery, lower cash inflow, and higher repayment pressure. The financial baseline failed to detect many cases early because cash-flow decline had not yet appeared in the statements. In contrast, the graph model connected penalties, emission-permit status, facility limits, supplier disruption, customer order decline, and loan pressure into warning paths. This result agrees with studies on environmental policy knowledge graphs, which show that structured environmental data can support risk analysis and decision-making [27,28]. For banks and industrial finance platforms, earlier warning can support credit-limit review, post-loan monitoring, and targeted field checks.

3.3 Compliance-Driven Risk Chain Analysis

Graph reasoning identified 6,240 compliance-driven risk chains and 2,080 enterprises affected by both environmental penalties and order decline. The most common chain was repeated penalty → late rectification → production limit → order decline → repayment pressure. This path shows how regulatory pressure can move from compliance events to operating problems and then to financing risk. Another common pattern was subsidy withdrawal after environmental non-compliance, followed by supplier payment delay and loan renewal difficulty. Fig.2 presents a typical compliance-finance risk chain linking environmental penalty, emission-permit status, production facility restriction, supplier disruption, customer order decline, and repayment outcome. This type of chain is more useful for bank review than a single risk score because it explains why a firm needs further investigation. Compared with ESG knowledge graph studies that mainly extract entities from news or policy texts, this study links compliance records with loan and transaction outcomes, making the reasoning results more suitable for industrial financing-risk prediction [29,30].

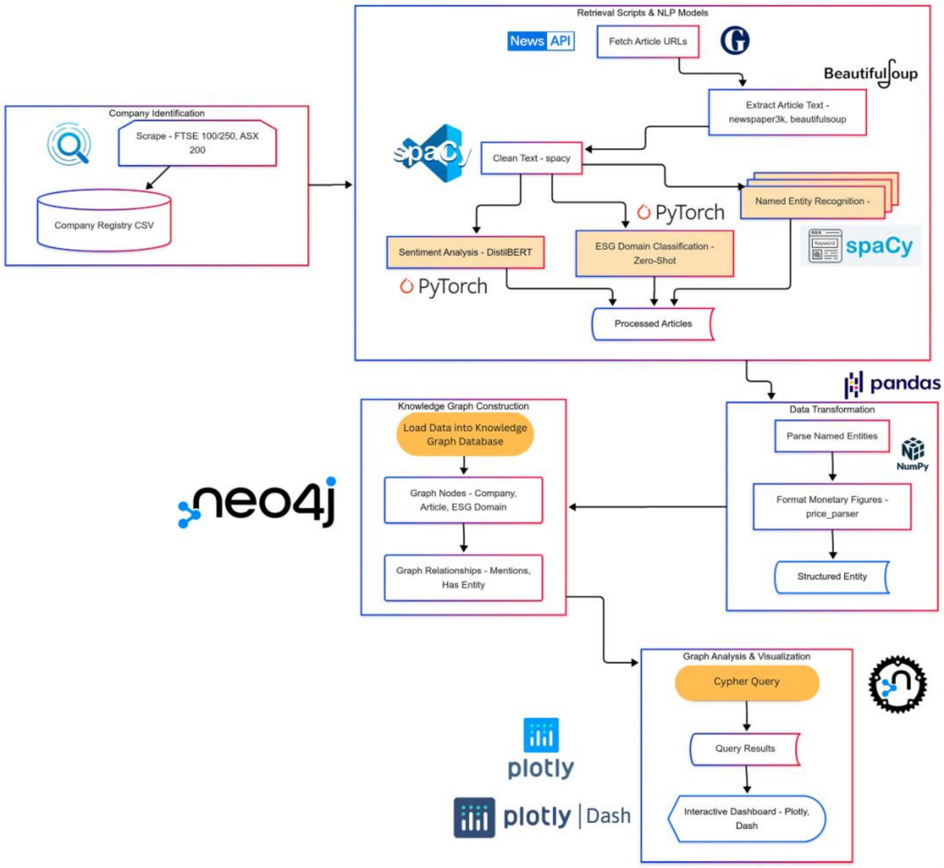

Fig. 2. Compliance-finance risk chain showing the links among penalties, production limits, order decline, and repayment pressure.

3.4 Practical Implications and Comparison with Prior Studies

The system grouped 15,700 raw enterprise alerts into 3,890 investigation cases through compliance-path aggregation. This result shows that graph reasoning can reduce repeated manual checks and help credit officers focus on firms with similar risk causes. Full monthly assessment was completed in 13.9 minutes, with a median inference latency of 47 ms per enterprise node. Therefore, the method can support routine financing-risk monitoring at regional or industrial-platform scale. Compared with earlier SME credit-risk studies, this work adds environmental inspection, penalty, emission-permit, facility, and pollution-control equipment data to the credit-risk framework. Compared with environmental compliance studies, it further examines whether compliance events are linked to financing outcomes. The results suggest that environmental compliance is not only a regulatory issue but also a financing-risk factor related to production continuity, order stability, supplier confidence, and repayment capacity. However, the model should be used as a decision-support tool. For high-risk cases, graph warnings should be checked together with field visits, bank statements, permit documents, rectification reports, and loan contract records.

4. Conclusion

This study developed an environmental compliance graph reasoning model to predict financing risk in industrial enterprises. By combining environmental penalties, emission permits, production limits, pollution-control equipment, supplier and customer networks, subsidies, loan contracts, and repayment outcomes into a knowledge graph, the model identified risk signals that traditional financial indicators often miss. The results showed higher prediction accuracy, a reduction in median warning time from 82 days to 30 days, detection of 6,240 compliance-driven risk chains, and identification of 2,080 firms affected by both penalties and order reduction. The system also consolidated 15,700 raw alerts into 3,890 investigation cases and completed monthly assessment of all enterprises in 13.9 minutes. The main contribution is the integration of graph-based representation learning with knowledge reasoning, which enables both early warning and interpretable risk analysis. These findings demonstrate that environmental compliance strongly influences financing risk and that relational information between firms, regulatory events, and operational constraints can improve credit evaluation. The method can support banks, industrial finance platforms, and regulators in loan screening, monitoring, and risk review. Limitations include the use of data from a single region and potential gaps in private enterprise records. Future studies should apply the model to other regions and industries, include more real-time operational data, and further analyze causal links in compliance-driven financing risk.

References

1. Wu, C., & Chen, H. (2025). Research on system service convergence architecture for AR/VR system.
2. Sipangkar, F. J., & Firmansyah, A. (2026). Environmental Performance, Green Accounting, Carbon Performance, And Firm Value: Does Board Meeting Frequency Matter?. Jurnal Riset Akuntansi Terpadu, 19(1), 97-128.
3. Du, Y. (2025). Research on Digital Quality Traceability System for Temperature-Controlled Supply Chain of Foreign Trade Wine Driven by Blockchain and IoT. Business and Social Sciences Proceedings, 4, 57-65.
4. Shao, W. (2026). Design and Implementation of an Open-Source Security Framework for Cloud Infrastructure. arXiv preprint arXiv:2604.03331.
5. Radhakrishnan, V. (2026). Role of Forensic Accounting in Detecting Flawed Control Measures in Financial Institutions Causing Regulatory Penalties (Doctoral dissertation, Trident University International).
6. Xu, D., Chen, H., & Gui, H. (2026, March). Unified Online Estimation Method for SOC, SOH, and Power Capacity Considering Safety Boundary Consistency in Battery Management Systems. In 2026 International Conference on Electrical, Control and Information Technology (ECITech) (pp. 133-137). IEEE.

7. 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.
8. Zhao, J., Fan, J., & Li, T. (2026). A Study on the Application of Industrial Document Understanding and Equipment Anomaly Reasoning in Discrete Manufacturing Operations and Maintenance. Available at SSRN 7259198.
9. Taheri Hosseinkhani, N. (2025). Global food security amid geopolitical tensions and climate risk: Trade governance and adaptive strategies. Available at SSRN 5448314.
10. Lin, T., Chen, A., & Spivack, Z. (2026). Research on Collaborative Piano Teaching Models and Teaching Resource Efficiency Optimization in Public Art Education Systems. Available at SSRN 6284818.
11. Qi, C., & Qiao, X. (2026, May). From Traditional Machine Learning to Large Language Models: The Evolution of Infrastructure from an Engineering Perspective. In 2026 6th International Symposium on Computer Technology and Information Science (ISCTIS) (pp. 503-506). IEEE.
12. Nsiah, J., Agyemang, A. O., & Bawuah, B. (2025). Advancing environmental performance in emerging economies: the roles of working capital management and capital expenditure. Cogent Business & Management, 12(1), 2566902.
13. Xu, T., Zhang, J., & Zhu, W. (2026). Reproducible Modeling Pipelines and Cross-Window Stability in Subprime Auto Loan Credit Risk Assessment. Available at SSRN 6893861.
14. Sinha, D., Tong, Y., Zepeda, M. A. F., Hanstad, G., Nelson, E. C., Theisen, C. O., ... & Gamm, D. M. (2023). Silica nanocapsules as a nonviral delivery platform for iPSC-RPE. Investigative Ophthalmology & Visual Science, 64(8), 774-774.
15. Eyinade, W., Ezeilo, O. J., & Ogundeji, I. A. (2025). Financial risk management strategies and their influence on organizational stability. Finance & Accounting Research Journal, 7(6), 229-251.
16. Xiong, W., & Guo, Y. (2026). A Study on the Phenomenon of Abnormal Load Concentration and Its Cross-system Transmission Characteristics in High-density Building Operations. Available at SSRN 7120661.
17. Hong, Z., Chen, H., & Xu, T. (2026). Long-Horizon Load Stability Testing and Degradation Forecasting for High-Concurrency Cloud Storage. Available at SSRN 7115859.
18. Wu, J., Wu, D., Zhang, J., & Peng, Y. (2026). Generative AI Feedback in Junior High School Artistic Creation Evaluation. Available at SSRN 7244838.
19. Vidal, M. E., Chudasama, Y., Huang, H., Purohit, D., & Torrente, M. (2025). Integrating knowledge graphs with symbolic ai: The path to interpretable hybrid ai systems in medicine. Journal of Web Semantics, 84, 100856.
20. Huang, J., Xu, T., Yin, J., & Yang, J. (2026). Data Quality Monitoring and Intelligent Scheduling Optimization in Financial Data Workflow Automation. Available at SSRN 7179338.

21. Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. Advances in Neural Information Processing Systems, 35, 12800-12812.

22. Tariq, N. (2025). A Geoai-Based Environmental Risk Assessment Framework For Chemical Exposure And Public Health Vulnerability In US Industrial Corridors. International Journal of Environmental Sciences, 11(25s), 2025.

23. Zhang, Z., Wang, J., Li, Z., Wang, Y., & Zheng, J. (2025). Anncoder: A mti-agent-based code generation and optimization model. Symmetry, 17(7), 1087.

24. Lin, N., Xu, Y., Yang, H., Zhang, G., Zhang, M., Wang, S., ... & Li, X. (2020). Dissociating the neural correlates of the sociality and plausibility effects in simple conceptual combination. Brain Structure and Function, 225(3), 995-1008.

25. Oreshile, S. A., Mahdzan, N. S., & Zainudin, R. (2025). Enterprise risk management quality and firm value: Evidence from corporate reputation risk theory. Risk Management, 27(1), 4.

26. You, S. (2026). Verifiable Audit Mechanisms in AI Compliance Automation: Scalability. Available at SSRN 6547458.

27. Khan, S. U., Zhao, Q., Wisal, M., Shah, K. A., & Shah, S. S. (2025). A data-driven Bayesian belief network influence diagram approach for socio-environmental risk assessment and mitigation in major ecosystem-and landscape-modifier projects. Sustainability, 17(8), 3537.

28. Wu, J., Zhang, J., Wu, D., & Peng, Y. (2026). Visual Transformation Mechanisms for Integrating Digital Cultural Heritage Resources into Junior High School Art Curricula.

29. Luchkov, K., & Velinova-Sokolova, N. (2026). Counterparty Anti-Money Laundering Risk: A Three-Factor Matrix Model for Assessment and Control. Risks, 14(7), 170.

30. 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.