ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

Knowledge Graph-Augmented RAG Approach: For Compliance Risk Identification and ESG Governance Assessment in Multinational Enterprises

Zhou Liwen

International Business School, Fuzhou University of International Studies and Trade, Fuzhou, China, 350202
87360379@qq.com

Abstract: *Compliance regulation in cross-border trade exhibits multidimensional characteristics—such as inconsistencies in documentation, fragmentation of compliance information, and difficulty in detecting risk propagation—making it challenging to accurately identify compliance-related matters and risks using solely traditional manual methods or Retrieval-Augmented Generation (RAG) technologies. To address these challenges, this paper proposes an enhanced risk identification method based on RAG. This approach leverages a domain-specific knowledge graph, integrates compliance management throughout the entire cross-border trade lifecycle, and incorporates it into an analytical framework. We construct an ontology model across three dimensions—risk themes, governance metrics, and factual entities—thereby building a cross-border trade compliance knowledge graph comprising 32 nodes and 54 semantic edges, which enables structured representation of rules and evidence. Furthermore, we optimize the conventional RAG workflow by employing techniques such as query rewriting, hybrid retrieval, result ranking, and graph prompting, thereby facilitating the extraction of high-confidence compliance-related items. We employ a two-hop weighting method to quantify the intensity of risk propagation and measure the exposure levels of various compliance-related elements, ultimately generating an interpretable corporate risk profile. The correlation patterns observed in our experimental results indicate the existence of latent compliance risks, particularly those associated with trade financing and anti-money laundering (AML) activities. These findings demonstrate that the proposed method is both feasible and practical, providing a technical solution to enhance compliance in cross-border trade operations.*

Keywords: *RAG; knowledge graph; cross-border trade; risk quantification; multinational corporations; international business; ESG governance; sustainable development*

1. Introduction

As globalization advances, trade regulations have become increasingly stringent. Cross-border trading enterprises must integrate export controls, customs classification, international sanctions, anti-money laundering measures, and the cross-border handling of relevant data into their compliance frameworks; any violation may result in penalties such as fines, revocation of licenses, or damage to brand reputation. The compliance documentation in cross-border trade varies widely in form and structure—encompassing customs documents, contracts, bills of lading, certificates of origin, etc.—and is characterized by fragmented data structures, highly specialized terminology, and complex interrelationships among these documents. Employing traditional manual review methods for such documentation is not only inefficient and error-prone but also difficult to scale, thereby failing to meet the demands of modern trade risk management [1][2][3].

In the broader context of international business, multinational enterprises (MNEs) operating across multiple jurisdictions face increasingly complex interactions among cross-border trade regulations, global supply chain governance, and growing expectations for sustainability. Beyond legal compliance, MNEs are also expected to integrate environmental, social, and governance (ESG) principles—such as responsible procurement, anti-bribery measures, cross-border data protection, and carbon accountability—into their global operations, treating trade compliance and sustainability as integral components of their corporate governance framework rather than as separate, independent functions.

Current mainstream compliance and risk management solutions typically rely on rule-based

ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

matching approaches or foundational large language model (RAG)-based approaches. Traditional methods—such as those employing keyword matching or rule engines—require manual definition of rules, can only perform coarse-grained filtering, and are unable to uncover implicit risk correlations across multiple documents. In contrast, approaches based on foundational RAG technology process text in isolation, without adequately considering domain-specific semantics or causal chains, which results in suboptimal risk interpretation outcomes. By comparison, knowledge graphs provide structured semantic support and possess multi-step reasoning capabilities, thereby effectively mitigating the limitations of large language models. Leveraging this advantage in compliance risk analysis can offer an innovative pathway for problem solving [4][5].

To address challenges in cross-border trade—such as inaccurate extraction of compliance information, difficulties in quantitative risk assessment, and a lack of interpretability in final outcomes—this paper proposes the following countermeasures: constructing a dedicated cross-border trade compliance knowledge graph; employing an enhanced multi-stage RAG approach for information extraction; and applying a weighted analysis method to examine risk propagation, thereby quantifying compliance risks [1][2][6]. The innovative contributions of this study are as follows:

(1) The structural compliance knowledge required for cross-border trade shall be designed using a three-tier architecture, taking into comprehensive consideration all relevant parties, validation standards, and business documents, while systematically and structurally encoding domain-specific knowledge.

(2) Integrate the multi-stage RAG optimization method into the workflow, combining hybrid retrieval techniques with graph-augmented prompting to accurately and reliably extract compliance-related information from diverse trade documents.

(3) We propose a quantification method employing a dual-jump weighting approach for quantifying risk transmission; we apply a hierarchical attenuation framework to address risk exposures across multiple themes; and we conduct a rankable and interpretable analysis of corporate compliance risks.

From the perspective of international trade and multinational corporate management, the aforementioned research findings can assist multinational corporations in aligning fragmented national regulations and addressing the heterogeneity of cross-border trade documentation, while also bridging compliance evidence at the transactional level with higher-level ESG governance indicators. The resulting risk profile thus serves not only as a legal compliance report but also as a decision-support tool for overseas investment, the formulation of export strategies, and the implementation of responsible supply chain management.

2. Core Algorithms

2.1 Improved RAG Information Extraction Algorithm

Traditional RAG approaches face challenges such as incomplete retrieval and suboptimal semantic alignment when processing specialized cross-border trade texts. This paper adopts a four-stage optimization framework: first, specialized trade terms are supplemented via query rewriting; second, a hybrid retrieval strategy combining BM25 keyword matching and DPR semantic retrieval is employed; third, a cross-encoder is applied for re-ranking and noise filtering; finally, knowledge graph-based semantic constraints are introduced to optimize the output of large language models and standardize entity confidence scores, thereby enhancing the accuracy and completeness of cross-border compliance information extraction for multinational enterprises.

2.2 Dual-Search Weighted Risk Identification Algorithm

Building upon a three-tier architecture composed of a compliance knowledge graph—consisting of "risk topics," "governance indicators", and "factual evidence" —this paper proposes a two-hop weighted risk quantification method. First, the confidence scores of multiple documents are aggregated to compute the knowledge score for each node; next, leveraging the hierarchical decay property of risk propagation, governance indicators are left unattenuated, whereas indirect evidence is attenuated using empirical coefficients. This process generates an exposure score for each risk topic, which are then sorted in descending order. Consequently, this method enables both quantitative and interpretable assessment of corporate compliance risks, while also aligning seamlessly with ESG governance indicators.

ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

3. System Design and Implementation

3.1 Overall System Architecture

This paper proposes a design methodology for an end-to-end cross-border trade compliance risk identification system. The system adopts a five-layer modular architecture, integrating data preprocessing, information extraction, knowledge modeling, risk quantification, and result visualization within a unified workflow. Built on the R programming language, the system leverages the igraph library for graph computation and the visNetwork package to enable interactive visualization; its overall architecture is lightweight and demonstrates strong practical feasibility. The data preprocessing module supports batch loading of diverse trade documents based on user requirements, performing text cleaning, deduplication, and format standardization to eliminate irrelevant noise and construct a standardized trade corpus. The enhanced RAG extraction module employs a multi-stage optimization process to identify compliance-relevant elements and assess their confidence levels. The knowledge graph module utilizes a domain ontology-driven structure to progressively integrate extracted information into the graph, while dynamically assigning knowledge scores to nodes and edges. The risk quantification module applies a two-hop weighting method to determine risk exposure levels and risk grades in a logical sequence. The visualization module is dedicated to displaying the graph structure, risk propagation paths, and various scoring indicators (with support for data export). Through the coordinated operation of these modules, raw textual data can be fully transformed into quantified risk outcomes [1][4][5][6].

At the architectural level, this platform is designed to address the compliance and ESG governance needs of international trade enterprises: it consolidates disparate trade documents, applies standardized processing using shared ontologies, and displays legal compliance scores alongside ESG-related risk exposures on a single dashboard. This lightweight, reusable tool can be adopted by trade companies and their multinational business units without relying on proprietary systems locked to a single vendor.

3.2 Building a Cross-Border Trade Compliance Knowledge Graph

Based on compliance regulations applicable to cross-border trade and actual business operational practices, this paper proposes a three-tier progressive domain ontology—comprising risk-related topics, governance indicators, and evidence entities—and integrates these components into a unified "Risk–Indicator–Evidence" reasoning framework. Finally, this paper also presents an example of a domain knowledge graph consisting of 32 nodes and 54 directed edges [3][4][6][7].

The risk thematic layer performs a foundational classification of top-level risks and enumerates eight of the most common compliance issues in cross-border trade: export control compliance, customs commodity classification and tariff compliance, (international) sanctions compliance, supply chain and third-party compliance, trade financing and anti-money laundering (AML) compliance, documentation and licensing compliance, anti-corruption (bribery) measure compliance, and cross-border data compliance; this layer provides a comprehensive overview of compliance risks across all stages of import and export activities. The edges in the diagram represent three types of structural relationships: 17 risk edges connect themes with indicators based on risk-related criteria; 25 support edges map indicators to evidence entities based on evidential logic; and 12 consistency edges facilitate horizontal comparisons between entities or verification against documentation. Each node is assigned a prior weight and a time-varying knowledge score; the category of each connecting edge and its dynamic connection score are also explicitly defined to strike a balance between professional judgment and flexible adaptation [7][8]. For further details, refer to Figure 1.

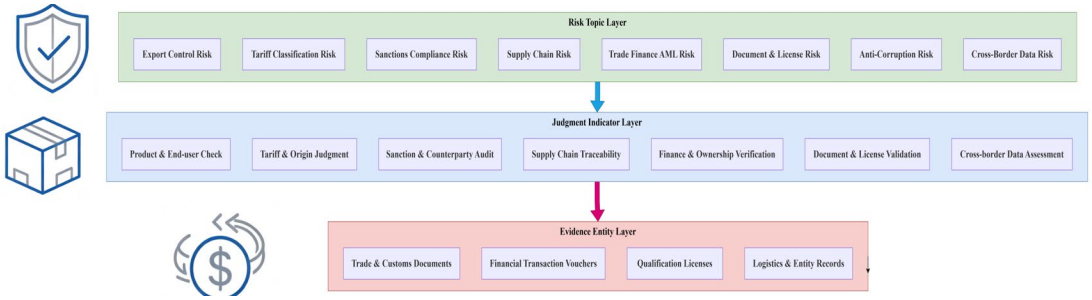


Fig. 1.Cross-Border Trade Compliance Knowledge Graph — Schematic Diagram of a Three-Level Ontology Architecture Design

ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

These eight risk categories can also be further interpreted from the perspectives of international commerce and ESG: export controls, sanctions, and anti-bribery measures correspond to the Governance (G) pillar; supply chain management and third-party due diligence align with the Social (S) pillar, encompassing responsible procurement and labor standards; cross-border data compliance and customs transparency, in turn, align with the broader Environmental and Sustainability (E) expectations. By aligning its ontology with this ESG framework, the knowledge graph can be utilized both for legal compliance and for non-financial information disclosure.

3.3 Implementation of RAG-based Information Extraction and Risk Quantification Using Image Enhancement Techniques

This paper leverages existing compliance knowledge graphs to address domain-specific information extraction tasks based on RAG, aiming to overcome the limitations of traditional models in domain matching and reasoning. The approach begins with multi-source heterogeneous trade documents; it expands the retrieval scope by appropriately modifying query statements, employs a hybrid sparse-dense retrieval strategy to achieve high recall rates, filters out irrelevant content using re-ranking rules, and then utilizes the knowledge graph to classify various textual elements as compliance items or risk points. Finally, this paper proposes a domain-specific prompt construction method [7].

In this paper, the term "large language model" refers to a method for information extraction that employs enhanced prompting techniques, generating triplets based on the (doc_id, node_id, confidence) criteria. By systematically classifying and organizing six known categories of cross-border trade documents into an input demonstration corpus for the model, we obtain 20 reliable entity records. The knowledge extracted from each node undergoes statistical analysis and classification processing; subsequently, edge weights are adjusted based on the co-occurrence relationships between relevant entities within the documents, thereby endowing the graph edges with dynamic semantics. The resulting knowledge graph can then be continuously updated and optimized according to business requirements, ensuring that its semantics precisely align with real-world application scenarios [8][9][10]. For further details, please refer to Table 1.

Table 1 Detailed Example Table of Entity Extraction Results from the Improved RAG Model

certificate of identification	Extract entity type	trust	Corresponding risk
1	customs declaration	0.92	Customs Duties and Commodity Classification Risks
2	ECCNcode	0.95	Export control compliance risk
3	Transportation Certificate	0.89	License compliance risk
4	commercial contract	0.93	Supply Chain Compliance Risks
5	Settlement period	0.88	Trade financing risks

By leveraging information from a novel knowledge graph to integrate multiple factors and applying the two-hop weighting method to the quantitative analysis of overall risk, this study employs an attenuation coefficient at the hierarchical level to reduce the weight of indirect evidence, thereby clearly distinguishing between direct compliance risks and risks arising from indirect correlations [10].

4. Experiments and Results Analysis

To evaluate the effectiveness of the framework proposed in this study in practical applications, we constructed a demonstration corpus comprising six categories of foundational documents (e.g., customs declaration forms, trade contracts and bills of lading, sanctions lists, etc.). In this research, we employed a native RAG model as the baseline model and subsequently applied an enhanced RAG framework to conduct comparative risk identification analyses. This study focuses on two core evaluation metrics: first, whether the model can successfully identify risk correlations based on a knowledge graph (i.e., whether it can uncover implicit compliance issues); and second, whether the ranked results of risk exposure align with cross-border trade practices. From a visualization perspective, the three-tier knowledge graph exhibits robust structural characteristics—comprising listed risks, governance elements, and supporting evidence—collectively forming a complete and traceable reasoning chain. In export control compliance analysis scenarios, this sub-graph clearly illustrates how risks propagate through verification checkpoints or business documents, thereby directly identifying the multi-level consequences arising from specific risks; compared to traditional localized scanning approaches, this methodology facilitates a more

ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

comprehensive understanding and interpretation of risks. For further details, please refer to Figures 2 and 3.

From the perspective of international trade, this study demonstrates that for the fragmented and multilingual evidence routinely handled by multinational corporations, traceable graphical reasoning is more suitable than document-level scanning. The identification of implicit risk correlations also highlights a core governance challenge in overseas operations: the implicit interconnections among trade financing, third-party involvement, and sanction risks, which could evolve into significant ESG and reputational risks.

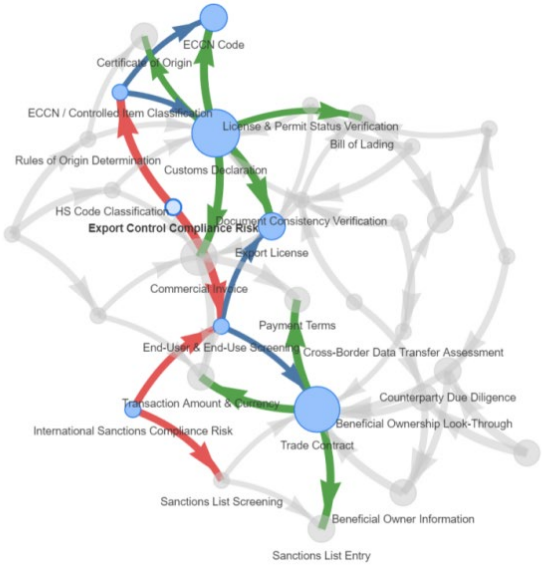


Fig. 2.Knowledge Graph and Schematic Diagram of Risk Propagation Path for Export Control Compliance Risks

Note: Red arrows indicate risk transmission paths; green arrows indicate evidence-supported relationships; blue nodes represent knowledge entities; this diagram illustrates the logical flow of how export control risks propagate to relevant indicators and knowledge entities.

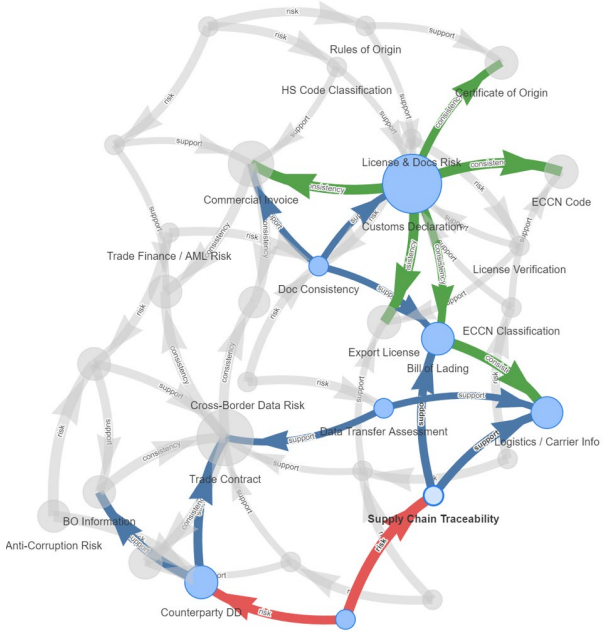


Fig. 3.Comprehensive Knowledge Graph for Cross-Border Trade Compliance

Note: Red arrows indicate risk propagation paths; green arrows indicate evidence-supported relationships; blue nodes represent knowledge entities. This visualization comprehensively illustrates the semantic associations among eight major risk themes, evaluation indicators, and evidence entities, as well as the structure of the risk propagation network.

ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

Among the listed risk categories, trade financing exhibits the highest risk score of 5.255 in terms of both risk score and prevalence, indicating that this category represents the most prominent risk domain under the current sample conditions; followed by cross-border data compliance (4.33) and export control compliance (3.54); whereas compliance matters related to international sanctions receive the lowest score (2.19), which aligns closely with the business context of the experimental dataset. These findings demonstrate that the proposed method in this study possesses robust capabilities to extract implicit compliance information from unstructured documents, enabling both reasonable quantification and precise identification of risks. For further details, please refer to Table 2.

Table 2 Risk exposure ranking sorted by compliance risk themes

social class	Compliance Risk Topic	Risk Exposure Score
1	Trade Finance and Anti-Money Laundering Risks	5.255
2	Cross-border data compliance risks	4.33
3	Export control compliance risk	3.54
4	Anti-corruption and bribery risks	3.515
5	Supply Chain and Third-Party Compliance Risks	3.475
6	Document and Licensing Compliance Risks	3.13
7	Customs Duties and Commodity Classification Risks	3.125
8	International sanctions compliance risks	2.19

5. Conclusion

In response to prevalent challenges in current cross-border trade practices—such as incomplete compliance information, overly simplified risk assessments, and a lack of interpretability in analytical results—this study proposes a compliance risk identification framework based on knowledge graphs and an enhanced RAG approach. The framework employs a three-tier logical structure: first, it maps cross-border trade-related knowledge onto a graph structure to enable structured analysis of domain-specific knowledge; second, it utilizes a multi-stage optimized RAG method to enhance the quality of compliance information extraction from heterogeneous trade texts; and third, it applies a two-hop weighting mechanism to quantitatively assess the relative risk intensity or severity of individual risk factors. From an practical application perspective, this framework effectively reveals implicit compliance correlations, accurately identifies key risk points for enterprises, and demonstrates clear interpretability and practical value. Future research will focus on three key directions: expanding the dataset using real-world cross-border trade data; establishing manually annotated gold standards for model benchmarking; and conducting quantitative performance evaluations of the models. Additionally, the study will compare various confidence-based methods to optimize the calculation mechanism for relevant risk weights, and integrate graph neural network techniques to enable dynamic risk assessment, thereby achieving more precise risk identification in complex operational scenarios.

Looking ahead, integrating this methodology with the ESG reporting frameworks and sustainable supply chain standards of multinational corporations will enable companies to align cross-border trade compliance with their global sustainable development strategies, thereby advancing the long-term objective of fostering responsible, transparent, and resilient international business practices.

References

[1] Pan Quan Tao, Zhang Lei, Liu Bailong, et al. A Graph-Based Traffic Prediction Model—Enhanced via Multi-Source Associative Knowledge Graphs [J]. Journal of Computer Science, 2026,53(8):245–256. DOI: 10.11896/jsjkx.250900014.

[2] Chang L. J., Lai C. M. FactualRAG: A RAG-based framework for fact-checking large language models—leveraging IFCN-certified sources [J]. Journal of Intelligent Information Systems, 2026,64(4):1687–1718. DOI: 10.1007/s10844-026-01041-y.

[3] Wang Ying, Wang Liya. Risk Identification and Countermeasures for Cross-Border Data Asset Trading in China [J]. International Trade, 2026(2):27–37.

*[4] Yu Zhi, Li Chunmei. A Tokenization-Based Cross-Border Trusted Data Space: Key Technologies, System Implementation, and Future Prospects [J]. *Journal of Telecommunications Science*, 2026.*

[5] Tao Rui, Zhang Sulan, Zhou Huiyuan. An image description generation model based on knowledge graphs and a global contextual attention mechanism [J]. Computer Engineering & Design, 2025, Vol. 46, No.5: 1297–1303.

ISSN 2616-7433 Vol. 8, Issue 4: 14-20, DOI: 10.25236/FSST.2026.080403

[6] Abo El-Enen M, Saad S, Nazmy T. A review of retrieval-enhanced generation (RAG) models for healthcare applications [J]. Neural Computing & Applications, 2025,37(33). DOI: 10.1007/s00521-025-11666-9.

*[7] Beckhauser W J, Fileto R. EchoRAG: A framework for enhancing language models through graph-RAG and contextual learning [J]. *Machine Learning*, 2025. DOI: 10.1007/s10994-025-06859-1.*

*[8] Huang Yining. Building a Resilient Regulatory Framework to Address Cyber Threats in Cross-Border Digital Trade: A Perspective Based on Systemic Risk Prevention [J]. *Research on Economic Issues*, 2025(4).*

[9] Lu Can, Gao Hui, Yang Jianlin. Research on the Interactive Relationship Between Value and Risk of Cross-Border Data Flows Based on Dynamic Simulation [J]. Journal of China Science and Technology Information, 2025, Vol. 44, No.12:1523–1538.

*[10] Mao Jie, Hao Zhenxing, Liu Hongzhong. The Impact of National Security on Risks Associated with Cross-Border Capital Flows: An Analysis Based on Disclosed Capital Flow Data [J]. *World Economic Research*, 2025(8).*