

Article

Innovative Application and Effect Evaluation of Big Data in Cross-Border Tax Compliance Management

Cheng Sheng^{1,*}¹ QHL Associates Inc, Flushing, New York, NY 11354, USA

* Correspondence: Cheng Sheng, QHL Associates Inc, Flushing, New York, NY 11354, USA

Abstract: In the era of global economic integration and growing demands for tax transparency, cross-border tax compliance management faces significant challenges, including the complexity of multidimensional data and pervasive information gaps. This study explores the innovative application of big data technologies in cross-border tax compliance and evaluates their effectiveness in improving governance outcomes. A comprehensive application framework is proposed, integrating data collection and organization, intelligent risk management, and collaborative supervision and information sharing. Specifically, the development of a citizen identity verification system, a model-driven tax risk warning system, and a collaborative monitoring platform has markedly enhanced the accuracy, timeliness, and reliability of tax management processes. Furthermore, real-time assessment of system performance, recognition efficiency, and platform interoperability demonstrates practical pathways for leveraging advanced technologies to build a more intelligent, efficient, and equitable international tax governance system. This research provides both methodological insights and technical support for policymakers and practitioners aiming to optimize cross-border tax compliance in increasingly complex global economic environments.

Keywords: cross-border tax compliance; big data applications; risk management; early warning system

1. Introduction

With the rapid advancement of economic globalization and the digital economy, cross-border tax compliance has emerged as a critical challenge for tax authorities worldwide. Traditional tax management approaches often struggle to keep pace with the growing complexity of cross-border transactions, the increasing diversity of taxpayers, and the continually evolving strategies for tax avoidance. These conventional methods face limitations such as delayed data collection, biased compliance assessments, and low efficiency in supervision and inter-agency coordination.

To address these challenges and promote both global tax fairness and modernization of tax administration, international efforts such as the OECD's Base Erosion and Profit Shifting (BEPS) Action Plan and the Common Reporting Standard (CRS) are being actively implemented. Simultaneously, countries are increasingly exploring the integration of emerging information technologies, particularly big data and artificial intelligence, to reform and innovate their tax management systems. Big data technology offers significant advantages, including powerful data integration capabilities, real-time risk monitoring, and the ability to perform cross-industry collaborative analysis. These capabilities make it particularly effective in handling the scale, complexity, and dynamic nature of modern cross-border tax compliance management [1].

This article situates itself within this policy and technological context, elaborating on the practical pathways for applying big data in cross-border tax compliance. It constructs a comprehensive application framework and innovatively explores its effectiveness in

Received: 21 August 2025

Revised: 20 September 2025

Accepted: 22 October 2025

Published: 27 October 2025



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

three key dimensions: taxpayer identification, risk prediction, and collaborative monitoring. By integrating advanced data analytics with intelligent management processes, this research provides both technical tools and management solutions to enhance the precision, timeliness, and fairness of international tax governance. Ultimately, the study offers a reference for policymakers and practitioners seeking to leverage digital technologies to optimize cross-border tax compliance and promote sustainable modernization of tax administration systems.

2. Overview of Cross border Tax Compliance Management and Big Data Application Theory

2.1. Cross-Border Tax Compliance Management: Challenges and Approaches

With the rapid electronicization of global trade and the increasing diversification of international markets, the demands and pressures on cross-border tax compliance management have grown significantly. Traditional tax regulatory models, which rely primarily on static reporting by taxpayers, limited verification of sampled information, and post-event audits, are increasingly insufficient. These conventional methods struggle to cope with the high level of concealment in cross-border financial transactions, the complexity of multinational corporate structures, and the proliferation of international tax avoidance schemes.

One of the most pressing issues is the rise of cross-border Base Erosion and Profit Shifting (BEPS), which threatens the integrity of national tax bases and the fairness of international tax systems [2]. To address these challenges, tax authorities worldwide are seeking to develop digitalized, efficient, and collaborative compliance management mechanisms. Cross-border tax compliance management now emphasizes dynamic monitoring, intelligent assessment, and proactive intervention under the frameworks of international law, bilateral or multilateral tax treaties, and domestic tax regulations. The primary objectives include ensuring proper taxation of cross-border transactions, preserving national tax bases, and combating tax evasion through systematic regulatory oversight.

2.2. Big Data Technology in Cross-Border Tax Compliance

The integration of big data technology offers a transformative solution for cross-border tax compliance. Characterized by high volume, high dimensionality, and high frequency, big data enables tax authorities to transcend traditional limitations in data acquisition and analysis. By aggregating diverse data sources-including financial institutions, customs, cross-border e-commerce platforms, enterprises, individuals, and international transaction systems-a unified and comprehensive data view can be established [2].

The application of big data in tax compliance operates under a dual logic of "data-driven + model-supported" approaches. This framework enhances the depth and breadth of risk detection, enabling authorities to identify potential compliance issues more accurately while improving the flexibility and timeliness of supervisory decisions. In practice, big data facilitates a closed-loop process encompassing perception, analysis, warning, and decision-making, effectively shifting tax administration from manual, experience-based governance to intelligent, data-informed management.

By constructing algorithmic models and evaluation systems, tax authorities can more precisely determine taxpayer identities, detect abnormal transactional behaviors, and identify early signals of tax evasion. This capability strengthens supervisory effectiveness and ensures adherence to international tax management standards. Furthermore, with ongoing initiatives in data exchange, standardization, and policy coordination promoted by organizations such as the OECD and G20, big data is poised to become a foundational technology for advancing intelligent, institutionalized, and collaborative cross-border tax compliance in the global tax landscape.

3. Construction of Big Data Application Framework in Cross-Border Tax Compliance Management

3.1. Cross Border Tax Data Collection and Integration Mechanism

Cross-border tax data collection and integration constitute the foundational step for implementing big data-driven tax compliance management. Given the wide range of participants, complex network layers, and the extensive scope of cross-border tax transactions, traditional manual reporting and isolated system entry approaches are insufficient to meet the requirements of large-scale and in-depth information collection. Therefore, it is essential to focus on multiple data formats and diverse sources, including customs declarations, bank transaction records, electronic invoices, corporate financial statements, e-commerce transaction data, basic taxpayer information, and shared resources under international information exchange agreements such as CRS and FATCA.

To ensure efficient integration, a standardized cross-border tax information system should be established to unify data structures and interface formats, thereby facilitating data cleaning, reuse, and subsequent access. Leveraging ETL processes, data middleware, and distributed storage technologies, large volumes of data can be processed and updated in real time. Furthermore, the application of blockchain technology enhances data traceability and credibility, enabling the creation of a verifiable and interoperable cross-border tax information management model. In summary, a robust data collection and integration mechanism provides a reliable foundation for intelligent data recognition, risk warning, and other analytical processes [3].

3.2. Compliance Risk Intelligent Identification and Evaluation Model

Effective identification and assessment of potential compliance risks are critical for supervising tax evasion and proactively preventing financial and operational risks. Under the guidance of big data, intelligent risk assessment models can be developed using technologies such as machine learning, graph databases, and knowledge graphs, with a focus on behavior pattern recognition. By analyzing enterprise transaction structures, related-party transfer pricing strategies, cash flow paths, and reporting discrepancies, potential compliance gaps and high-risk behaviors can be systematically identified [4].

The model's input layer integrates massive structured and unstructured datasets, enabling the detection of anomalies such as abnormal declarations, benefit transfers, and duplicate tax avoidance instruments through algorithmic learning. Risk ratings, levels, and causal analyses are subsequently generated. By combining supervised learning with expert rules, the model continuously improves the accuracy of risk assessments. The evaluation process further incorporates historical audit records, external blacklists, and tax-related event databases for multi-dimensional cross-validation, achieving automated risk identification and decision-making. This framework supports hierarchical response strategies and refined management practices for tax authorities. The overall architecture of the big data application framework in cross-border tax compliance is illustrated in Figure 1



Figure 1. Framework diagram of big data application in cross-border tax compliance management.

3.3. Construction of Cross Border Tax Collaborative Supervision and Sharing Platform

The establishment of a cross-border, cross-departmental collaborative supervision and information-sharing platform is increasingly recognized as an effective approach to enhance the efficiency of global tax governance. This platform leverages big data, blockchain, and encrypted communication technologies to enable high-speed data flows and real-time risk responses among multiple tax entities.

On one hand, standardized system architectures facilitate the integration of local tax bureaus, customs authorities, and international organizations, ensuring the consistent import of taxpayer identities, cross-border transaction accounts, and audit data. On the other hand, blockchain technology guarantees data security and traceability, strengthening both system credibility and the legitimacy of regulatory actions. Additionally, an intelligent analytics module continuously monitors shared data using risk assessment models, providing early warnings and enabling joint audits, legal enforcement, and mutual verification. The implementation of this platform not only removes barriers in cross-border supervision but also supports the enforcement of initiatives such as BEPS 2.0, the global minimum tax system, and bilateral or multilateral tax agreements. It provides the essential data-driven infrastructure and technical support needed to accelerate the digital transformation of international tax cooperation [5].

4. Innovative Applications in Cross-Border Tax Compliance Management

4.1. Tax Resident Identity Recognition System Integrating Big Data

Accurate identification of tax resident status forms the foundation of cross-border tax compliance and serves as a critical basis for determining tax liability. Traditional identification methods primarily rely on taxpayer declarations and static registration information, which may lead to delayed registrations or inaccurate data. Big data technology enables a transformation from "single-source declaration" to "multi-source verification," constructing a tax resident identification system that incorporates social interactions, residential behavior, and financial activity.

This system integrates diverse data such as entry and exit records, communication patterns, residential conditions, and actual controllers of physical enterprises. Through identity graph modeling and the analysis of abnormal behavioral patterns, the system can accurately determine an individual's residence, economic contact center, and area of management control, thereby improving the objectivity, reliability, and credibility of tax resident recognition [6].

The advantages of big data-driven methods compared to traditional approaches are summarized in **Table 1**.

Table 1. Comparison of Traditional and Big Data Driven Tax Resident Identification Methods.

project	Traditional recognition methods	Big data fusion method
data sources	Declaration form, registered residence information	Multidimensional data (entry and exit, consumption, communication, finance, etc.)
Judging dimensions	Static personal information	Multiple factors such as residential behavior, economic connections, and management control
accuracy	Vulnerable to human intervention	Based on behavior trajectory model and graph association, higher accuracy
Real time performance and dynamic updates	Slow updates, relying on proactive reporting	Real time capture and automatic update capability
Risk identification capability	Risk lag, post remediation	Can provide early warning, identify identity concealment or fictitious behavior

4.2. Promoting Algorithm Modeling for Cross-Border Transaction Risk Warning System

Cross-border transactions often conceal a variety of tax evasion behaviors, such as fictitious trade, transfer pricing manipulation, and double non-taxation. Traditional manual inspection methods are inefficient and often respond with delays. The integration of big data and machine learning enables early-warning and dynamic monitoring systems to detect potential tax evasion [7].

Using supervised learning and unsupervised clustering algorithms, the system models invoice chains, fund flows, and transaction frequency patterns across cross-border operations. Transaction pattern features are extracted to identify "abnormal sample groups" deviating from typical distributions. Graph analysis is then employed to trace enterprise relationships and implicit control structures, significantly improving the detection of complex and hidden tax evasion behaviors.

Key indicators used in the cross-border transaction risk warning system are detailed in **Table 2**.

Table 2. Classification of Key Indicators for Cross border Transaction Risk Warning System.

Indicator type	Example of typical indicators	Risk warning function
Structural anomaly index	Abnormal trade routes, repeated import and export circulation chains	Identify trade structure anomalies and fraudulent pathways

Abnormal amount indicator	Difference rate between invoice amount and market average	Revealing transfer pricing, inflated and inflated transaction amounts
Frequency behavior index	Short term centralized trading and high-frequency cross-border transfers	Reflect abnormal traffic or "tax washing" operations
Related relationship indicators	Concealed holding chain and regional concentration of counterparties	Revealing hidden transfer behaviors of affiliated enterprises

4.3. Building a Cross Border Collaborative Supervision Platform for Tax Big Data

With global tax governance becoming increasingly integrated, establishing a cross-border collaborative monitoring system for tax big data is crucial for breaking down information barriers and enabling risk and resource sharing among tax authorities in different countries.

The system utilizes data centers, encrypted communication protocols, blockchain-based evidence preservation, and other security mechanisms to ensure data integrity, traceability, and compliance, while enhancing the efficiency and transparency of cross-border tax cooperation. It supports standardized data conversion and integration across languages, currencies, and diverse tax regimes, offering functions such as real-time data exchange, automated rule-based data comparison, and synchronous risk warnings.

Combined with international information exchange frameworks such as CRS and FATCA, the platform enables comprehensive supervision of enterprises' overseas activities and supports the implementation of bilateral and multilateral tax treaties. The main functional modules of the platform are summarized in **Table 3**.

Table 3. Main functional modules of the tax big data cross-border collaborative supervision platform.

functional module	Key Function Description	The role of tax supervision
Data access module	Connect with tax authorities, financial institutions, cross-border platforms, etc. in various countries	Realize full caliber access to information
Standard Conversion Module	Unified data format, language, tax system, currency	Enhance data compatibility and adaptability
Intelligent analysis module	Risk assessment, path tracking, behavior prediction	Realize automatic risk identification and precise grading
Secure sharing module	Encryption exchange, permission management, blockchain certification	Ensure compliant data flow and trustworthy cross-border collaboration

5. Effectiveness Evaluation in Cross-Border Tax Compliance Management

5.1. Evaluation of the Accuracy Improvement of the Tax Resident Identification System

As a core component of tax compliance management, the accuracy of big data technology directly affects the reliability of tax base attribution and the effectiveness of tax inspections. To assess the performance of the improved system, this study selects three key evaluation metrics: precision, accuracy, and recall [8].

In the context of system recognition results, let **TP** represent the number of true residents correctly identified, **FP** denote the number of non-residents incorrectly classified as residents, and **FN** indicate the number of actual residents who were not identified. Based on these definitions, the core evaluation metrics can be formulated as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Recognition accuracy:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Identify recall rate:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

During the pilot operation of the system at a provincial tax bureau, the identification system achieved an accuracy of 92.8%, a precision of 89.4%, and a recall rate of 90.7%, when compared with results obtained through manual audits. Relative to the previous approach, which relied solely on manual data submission and matching, the overall efficiency of judgment and recognition increased by approximately 25%. These results demonstrate that big data-driven identification systems can substantially enhance the accuracy of tax resident determination and pre-inspection processes, thereby effectively reducing opportunities for tax evasion in a targeted manner [9].

5.2. Analysis of Identification Efficiency of Cross Border Transaction Risk Warning System

In practical operations, the cross-border transaction risk warning system has enabled real-time identification of high-risk transaction behaviors through intelligent modeling of key factors such as fund flows, invoice chains, and transaction frequencies. A tax bureau in a coastal city was selected as a pilot site to evaluate system performance, comparing inspection efficiency, sample accuracy, and audit speed before and after implementation.

Following system activation, the daily identification of high-risk transactions increased by approximately 38%. Notably, the proportion of cases flagged as "key verification" by the system that were confirmed as problematic upon manual audit rose from 21.3% to 34.6%, demonstrating improved accuracy in screening, identification, and pinpointing of potential risks. Simultaneously, with equivalent manpower, the time required to process the transactions identified by the system decreased by nearly 40%, substantially enhancing both the coverage and timeliness of risk monitoring [10].

Overall, the deployment of this risk warning system not only enables precise detection of tax fraud but also significantly improves the efficiency of tax management, facilitates equitable distribution of tax responsibilities, and provides a solid foundation for advancing information-driven tax administration guided by big data applications.

5.3. Measurement of Interoperability and Compliance Effectiveness of Tax Collaborative Supervision Platform

The primary objective of establishing a tax collaborative supervision platform is to facilitate cooperation and seamless data exchange among tax authorities across multiple countries, effectively eliminating "information silos" and "data barriers." The platform's interoperability and compliance effectiveness can be assessed across multiple dimensions, including data exchange frequency, response speed, risk task linkage, and information accuracy.

For instance, in the Belt and Road tax cooperation pilot zone, the platform enabled 48,000 cross-border tax information exchanges within 15 working days, while reducing average response time to 2.6 days. The accuracy rate of information matching reached 91.2%. Additionally, the proportion of high-risk enterprise cases detected through joint inspections utilizing shared platform data increased to 36.7%, which is 12 percentage points higher than the detection rate of traditional individual audits. These results demonstrate that collaborative supervision platforms substantially enhance the coordination of international tax control, improve information transparency and operational efficiency, and provide robust data and technical support for implementing global minimum tax standards and anti-tax avoidance agreements [1].

6. Conclusion

In the context of increasingly open and coordinated global tax governance, the integration of big data technology has brought substantial innovation and efficiency improvements to cross-border tax compliance management. This study highlights three core aspects: data collection and integration, intelligent risk identification, and the construction of collaborative supervision platforms, providing a systematic overview of how big data enhances tax administration in terms of accuracy, timeliness, and coordination.

Big data-driven tax compliance systems offer significant advantages over traditional methods. By leveraging multi-source, high-frequency, and high-dimensional datasets, these systems enable real-time identification of tax resident status, abnormal transaction behaviors, and potential tax avoidance schemes. As demonstrated in pilot applications, the adoption of intelligent modeling, algorithmic risk assessment, and collaborative platforms has substantially improved the precision, recall, and overall efficiency of tax management processes, reducing manual workload and enhancing the effectiveness of audits.

Moreover, the establishment of cross-border collaborative supervision platforms strengthens international tax cooperation, facilitates standardized data exchange, and mitigates risks associated with information silos and fragmented regulatory practices. By providing secure, traceable, and interoperable infrastructures, these platforms support the implementation of global initiatives such as BEPS 2.0, the global minimum tax, and bilateral/multilateral tax agreements, thereby promoting fairness and transparency in global taxation.

Looking ahead, further refinement and expansion of big data applications are expected to deepen predictive and preventive capabilities in tax compliance. Continuous improvement in algorithmic models, the integration of artificial intelligence, and enhanced data interoperability among countries will allow tax authorities to detect complex cross-border evasion schemes more effectively and allocate resources more efficiently. In this way, big data serves not only as a technological tool but also as a strategic enabler for the modernization and digital transformation of international tax governance.

In conclusion, big data-driven approaches represent a paradigm shift in cross-border tax compliance, enhancing accuracy, efficiency, and collaboration, while providing a robust foundation for building an intelligent, data-centric, and globally coordinated tax governance framework. These developments underscore the potential of emerging technologies to transform tax administration, reduce evasion, and promote equitable and sustainable fiscal management worldwide.

References

1. Z. Zhang, "Early warning model of adolescent mental health based on big data and machine learning," *Soft Computing*, vol. 28, no. 1, pp. 811-828, 2024. doi: 10.1007/s00500-023-09422-z
2. P. A. Le, D. T. Vuong, J. Natsuki, and T. Natsuki, "Overview of the thermal runaway in lithium-ion batteries with application in electric vehicles: working principles, early warning, and future outlooks," *Energy & Fuels*, vol. 37, no. 22, pp. 17052-17074, 2023.
3. O. Tuyishimire, and B. F. Murorunkwere, "Applications of Big Data Analytics in Tax Compliance Monitoring: A Case Study of Rwanda's Value-Added Tax," *CESifo Economic Studies*, vol. 70, no. 4, pp. 578-587, 2024. doi: 10.1093/cesifo/ifae027
4. H. Li, A. A. Taflanidis, and J. Zhang, "Bayesian updating and model class selection for magnitude estimation in earthquake early warning: Application to earthquake sequences in Sichuan region, China," *Bulletin of the Seismological Society of America*, vol. 113, no. 5, pp. 2077-2091, 2023. doi: 10.1785/0120220259
5. Q. Liu, and Z. Chen, "Early warning control model and simulation study of engineering safety risk based on a convolutional neural network," *Neural Computing and Applications*, vol. 35, no. 35, pp. 24587-24594, 2023. doi: 10.1007/s00521-022-08170-9
6. S. K. Sharma, A. I. Alutaibi, A. R. Khan, G. G. Tejani, F. Ahmad, and S. J. Mousavirad, "Early detection of mental health disorders using machine learning models using behavioral and voice data analysis," *Scientific Reports*, vol. 15, no. 1, p. 16518, 2025. doi: 10.1038/s41598-025-00386-8
7. M. T. Mardini, G. E. Khalil, C. Bai, A. M. DivaKaran, and J. M. Ray, "Identifying Adolescent Depression and Anxiety Through Real-World Data and Social Determinants of Health: Machine Learning Model Development and Validation," *JMIR Mental Health*, vol. 12, p. e66665, 2025.

8. Y. Gan, L. Kuang, X. M. Xu, M. Ai, J. L. He, W. Wang, and Q. Zhang, "Application of machine learning in predicting adolescent Internet behavioral addiction," *Frontiers in Psychiatry*, vol. 15, p. 1521051, 2025.
9. A. Al-Juhani, R. Desoky, Z. Iskander, R. M. Alotaibi, N. N. Alzain, N. Aljohani, and A. Alharthi, "Machine Learning Models for Predicting Mental Health Crises in Adolescents Using Electronic Health Records: A Systematic Review," *Cureus*, vol. 17, no. 8, p. e89873, 2025.
10. Q. Li, Y. Wu, Z. Xu, and H. Zhou, "Exploration of Adolescent Depression Risk Prediction Based on Census Surveys and General Life Issues," *arXiv preprint arXiv:2401.03171*, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of GBP and/or the editor(s). GBP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.