International Business & Economics Studies
ISSN 2640-9852 (Print) ISSN 2640-9860 (Online)
Vol. 7, No. 2, 2025
www.scholink.org/ojs/index.php/ibes

Original Paper

Research on the Automation Path of Revenue Audit for E-commerce Platforms Based on RPA—Take Pinduoduo's "10 Billion Subsidy" Project as an Example

Ying Zhao[1]

[1] Dalian University, Dalian City, Liaoning Province, China

Received: February 21, 2025 Accepted: April 03, 2025 Online Published: April 21, 2025

doi:10.22158/ibes.v7n2p190 URL: http://dx.doi.org/10.22158/ibes.v7n2p190

Abstract

This study takes the "Billion Subsidy" project of Pinduoduo as an example to address issues in e-commerce auditing such as complex rules, large data volumes, and low efficiency in multi-system collaboration. By leveraging RPA technology, a flexible rule configuration tool, full-process data tracking, and intelligent risk screening mechanisms are constructed. An automated audit solution is developed from order generation to fund settlement. During implementation, business personnel can directly adjust rules through visual tools without technical intervention, and intelligent matching technology is used to resolve cross-system data format differences, establishing a human-machine division of labor mechanism. The practice of the "Billion Subsidy" demonstrates that this solution can significantly reduce audit time while enhancing risk identification capabilities and shifting audit focus from post-event inspections to real-time warnings. The experience gained from this study provides a reference for optimizing business rules and can be extended to scenarios such as risk prediction, offering practical pathways for the digital transformation of e-commerce auditing.

Keywords

RPA technology, e-commerce audit, automated audit, human-machine collaboration

1. Introduction

(1) Theoretical framework innovation of digital transformation of e-commerce audit

Traditional audit models face three core contradictions in the rapid development of e-commerce: the structural conflict between frequent business rule iterations and lagging audit standard updates, the efficiency gap between massive data scale and sampling verification efficiency, and the objective limitations of manual verification accuracy due to multi-system collaboration requirements. RPA

technology, through building a dynamic rule engine (pre-configured basic logic units for agile configuration), full data tracking (covering the entire order lifecycle), and intelligent collaborative validation (automatic cross-system data matching), drives the acquisition of audit evidence from probabilistic verification to deterministic analysis, forming a new paradigm of "digital mirror auditing." This transformation restructures the audit control system, upgrading risk identification from threshold triggers to pattern predictions, extending monitoring granularity from transaction levels to user behavior dimensions, and achieving a fundamental breakthrough in audit timeliness from delayed inspections to real-time warnings.

(2) Evolution and theoretical challenges of human-machine collaborative driven cognitive system

Technology application fosters a new audit cognitive framework: By transforming expert experience into iterative rule algorithms and leveraging the complementary strengths of machine processing structured data with human analysis of unstructured clues, a closed-loop mechanism of "data-insight-decision" is established. This collaborative model shifts audit functions from error tracing to risk prevention, forming a continuously optimized risk management ecosystem. The current theoretical system needs to address three major propositions: establishing an elastic rule framework to balance compliance rigidity with business innovation, setting legal standards for the evidentiary value of machine-audited evidence, and building ethical coordination mechanisms for algorithm transparency and audit independence, providing new coordinates for the development of audit theory in the digital economy era.

2. The Theoretical Construction and Practical Verification of RPA Technology in E-commerce Revenue Audit

RPA (Robotic Process Automation) is a key enabling technology for digital transformation. Its technical core lies in simulating the interaction behavior between human and digital system through software robot (Zhang, 2025) In the e-commerce revenue audit scenario, this technology system consists of three core components:

(1) Process analysis layer: Based on DOM tree parsing and computer vision technology (Deng, Zhang, Dong et al., 2025), it can accurately locate dynamic page elements (such as Pinduoduo subsidy activity pop-up), and can adapt to the technical challenges of 3-5 iterations and updates of the interface every month;

(2) Rule implementation layer: built-in subsidy rule engine containing 78 decision nodes, supporting the automatic processing of complex business logic such as "full reduction gradient calculation" and "cross-platform order mapping";

(3) Abnormal disposal layer: The integrated optical character recognition (OCR) and API interface monitoring (Qin, Lu, Yu et al., 2025) realize self-healing processing for 12 kinds of abnormal scenarios such as verification code failure and system response timeout (see Table 1).

Table 1. Comparison between RPA and Traditional Audit Methods

Compare dimensions	traditional audit	RPA technology
Deployment mode	The system code needs to be modified, and there are compliance risks	Non-intrusive deployment, simulating manual clicks
data-handling capacity	Only structured data is supported	Multi-source heterogeneous data (API + image recognition)
expansibility	Fixed processing capacity, unable to cope with traffic surge	hour；hr.
error rate	12.4%	0.35%
Audit cycle	72 hours	2.5 hours

Compared to traditional automation solutions, RPA demonstrates three significant advantages: First, its non-intrusive deployment feature allows data scraping through simulated human clicks without modifying the underlying code of e-commerce platforms, thus avoiding compliance risks associated with system modifications (Liu & Zeng, 2025); Second, its capability to handle multi-source heterogeneous data, which involves coordinating API interfaces (for structured transaction data) and image recognition technology (for parsing unstructured promotional pop-ups) to achieve comprehensive audit data collection (Zhang, 2025); Third, its elastic scaling mechanism, which, when the "billion subsidy" campaign triggered a 300% surge in order volume, increased system throughput from 5,000 transactions per hour to 15,000 transactions per hour using dynamic scaling techniques for robot clusters (Yang, Yang & Zha, 2024) (see Table 2).

By integrating a dynamic rule engine with intelligent comparison, the system achieves agile response to frequent business changes (Keast, 2025) A layered processing architecture automates standardized verification tasks while maintaining flexibility for human intervention, forming a closed-loop management system for risk classification warnings and handling (Mao, Xu, Jiang et al., 2024). The core innovation lies in establishing a linkage mechanism between rule configuration and execution feedback, ensuring strategy iteration efficiency through visual tools and enhancing risk identification accuracy through cross-validation from multiple sources. Ultimately, this forms a human-machine collaborative work model adaptable to various scenarios, providing an expandable implementation framework for transforming business processes from passive verification to proactive prevention (Success story about Robotic Process Automation (RPA) in Germany, 2025).

Table 2. The Improvement of Various Indicators after the Implementation of RPA

index	Implementation phase	Implementation	The extent of the increase
Order verification speed	120 m/s	5,800 lines/hour	47.3 times
error rate	12.4%	0.35%	Decreased by 97.2%
Audit cycle	72hours	2.5hours	96.5% shorter
Cross-platform order correlation success rate	83%	99.3%	16.3% increase
Detection rate of abnormal transactions	65%	96.8%	31.8% increase

3. Analysis of the Advantages of Introducing RPA in the Revenue Audit of Pinduoduo's "10 Billion Subsidy" Project

(1) Dynamically adapt to complex business rules

The e-commerce promotion rule system includes many dynamic elements, such as product category grading, user behavior constraints, and time slot strategy combinations. The frequency of its adjustments is often closely related to the marketing rhythm. Traditional audit models still rely on manual parsing of scattered rule documents, which can lead to issues like chaotic version management and misinterpretation of guidelines (Møller & Sauer, 2025). When multiple promotional activities run in parallel, the maintenance cost of paper documents can increase exponentially, often causing the audit baseline to diverge from current strategies, potentially leading to discrepancies in calculation standards (Pang, Song, Jiang et al., 2025).

(2) Process billions of data assets in full

Traditional auditing is like fishing with a net; it can only sample part of the data, inevitably missing risks when dealing with massive transactions. Now, through multi-system collaboration, the system automatically captures and organizes order, payment, and logistics information from different platforms, piecing together scattered data accurately like a puzzle. The system can automatically identify differences in data formats across platforms and make intelligent adjustments, completely breaking down information silos. From order creation to delivery, every step is traceable, turning the error-prone manual verification process into a precise automated assembly line.

The innovative integration of business rule libraries and machine learning models at the risk identification layer, along with the construction of a multi-dimensional risk profiling system, enables the system to automatically label the characteristics of abnormal transactions through a pattern recognition engine. It can implement tiered warnings based on risk levels—from automatically intercepting clear violations to pushing clues for human review. In conjunction with an interactive

analysis panel, auditors can quickly penetrate and view original business documents and relevant links, upgrading traditional spot-checking methods to a mechanism of "initial screening by machines + expert judgment," achieving a significant transformation from point detection to comprehensive monitoring in risk prevention.

(3) Build a real-time risk control response system

Traditional auditing is limited by batch processing mechanisms, leading to minute-level response delays when dealing with real-time transaction scenarios. The RPA system connects directly to the payment clearing center via an API gateway and uses real-time scanning technology to detect multiple risk dimensions such as user access heat maps and fund flow topologies at the millisecond level. When the system detects that the frequency of subsidy claims exceeds preset thresholds or that payment routes involve cross-regional jumps, it automatically triggers a two-factor authentication engine to complete risk interception actions in less than one second.

Risk management employs a four-tier circuit breaker mechanism, ranging from automatic correction of field misalignments and other routine anomalies to initiating manual review of suspicious fund loops. Each transaction generates an operation trace with timestamp watermarks, supporting bidirectional traceability verification. By constructing an enhanced loop of "machine pre-screening-expert diagnosis-strategy feedback," it achieves a throughput capacity of over 3,000 orders per second while ensuring high accuracy in identifying complex cash-out patterns. This architecture transforms the risk prevention system from lagging investigation to real-time monitoring, forming a comprehensive digital sentinel defense across the entire chain.

4. Implementation Path of RPA Technology for Income Audit of Pinduoduo's "10 Billion Subsidy" Project

(1) Systematic implementation stage

During the deployment phase, the "main trunk first-branch iteration" implementation strategy was adopted. First, a data transmission main channel based on HTTPS two-way authentication was established, integrating core systems such as the order center, payment gateway, and marketing platform. To address the semantic ambiguity issues in coupon rule texts, a regular expression engine + AST syntax tree converter was developed to transform complex marketing logic, such as "cross-store full reduction" and "deposit expansion," into executable decision nodes. A specially designed SM4 dynamic rotation encryption mechanism, combined with field-level data desensitization, ensures the integrity and confidentiality of audit data during cross-system data flow.

At the architecture level, a microservices design is adopted, breaking down into independent components such as data bus services, rule compilers, and risk decision engines. The data bus uses a read-write separation architecture, leveraging message queues to handle peak and off-peak loads of hundreds of millions of daily orders; the rule compiler supports hot-swappable policy package loading, ensuring new audit rules take effect in seconds during major promotions. The reserved extension

interface has been successfully integrated with the user profiling system, enabling real-time fusion computation of "subsidy sensitivity" tags. This highly cohesive and loosely coupled design pattern ensures the system maintains extremely high availability when facing sudden traffic surges from live streaming sales.

(2) Intelligent verification mechanism embedded

Based on the automation of basic processes, a three-tier verification mechanism is constructed to enhance audit reliability. This includes data integrity checks and real-time monitoring of data transmission, ensuring no loss or tampering through consistency validation. Business logic verification: an intelligent verification module is deployed to conduct multi-dimensional logical validation of subsidy calculation paths, ensuring that amount calculations comply with rules. Abnormal transaction screening: an intelligent detection model is established to identify potential abnormal transactions through pattern analysis.

(3) Construction of dynamic risk control system

This system automatically adjusts the inspection intensity based on the popularity of promotional activities. By learning from past audit experiences, the system assigns scores to different risk levels and automatically increases the sampling ratio during major promotions like Double Eleven. It has prepared over 20 common issue response plans in advance, such as immediately initiating the corresponding handling procedure when order volumes suddenly surge, automatically generating a to-do list.

(4) Optimization of human-machine collaboration mechanism

Using a "machine + human" relay work model, tasks are assigned like sorting packages on an assembly line. Basic tasks such as verifying order amounts and checking if coupons have expired are quickly handled by automated programs; for complex situations suspected of "scalping," the system automatically packages user behavior records and correlates them with order information. The intelligent audit knowledge base constructs a self-evolving risk feature mechanism through real-time linkage between machine validation and manual review. The system automatically captures multiple characteristic dimensions, including equipment association graphs and operation path trajectories, marked when new risks are handled manually, and dynamically generates incremental rule packages to inject into the screening engine. When variant risks similar to "cross-store order bundling refunds" are identified, the system triggers a dual-channel verification: automated components perform most standardized feature matches, while the remaining few suspicious cases are automatically added to the human judgment queue, accompanied by historical case handling records as references. This spiral optimization mechanism of "feature extraction-rule encapsulation-cross-validation" enables the risk model to automatically iterate through three versions each week, preserving the discretionary space for human judgment in gray areas while ensuring that the timeliness of experience solidifies ahead of the innovation speed of black market models.

(5) Continuous iterative mechanism design

Establish an improvement system based on the PDCA cycle, setting up a matrix of process health

assessment indicators that covers 12 core metrics, including task completion rate, anomaly capture rate, and manual review pass rate. Develop a version control module to support gray release; new functional modules must pass sandbox environment stress testing and logical validation tests before going live. Set up a cross-departmental collaborative improvement team to regularly collect optimization suggestions from multiple perspectives such as finance, operations, and technology, ensuring that the audit system evolves in sync with business development.

5. Implementation Points and Challenges

(1) Implementation points

1) Business rule analysis and process design

Full-chain business analysis involves a specialized team composed of audit, business, and technical personnel. They delve into the platform's promotional activity rules, tracing the entire process from order placement to fund settlement. The focus is on resolving complex issues such as cross-system discount stacking and activity priority conflicts, creating an audit flowchart that covers the entire lifecycle.

Dynamic rule adaptation mechanism, establish a rule change monitoring system, capture platform policy updates in real time (such as new promotion gameplay), automatically trigger audit model optimization, and form a fast response chain from rule adjustment to process update.

2) Security system integration scheme

The data security system adopts a triple protection design of "transmission encryption-storage isolation-permission cutover." Based on the separation of three powers model, a permission control matrix is constructed, allowing auditors to obtain field-level data access permissions and rule engineers to perform policy adjustments. A specially designed hardware encryption module implements secondary encryption for sensitive fields, ensuring that even if there is an unauthorized access, the core subsidy distribution logic remains protected at the physical level.

3) Human-machine collaboration optimization mechanism

The system adopts a dual-checking model of "machine self-inspection + manual spot checks," forming a virtuous cycle of continuous optimization. In daily operations, the system automatically identifies two types of issues: simple problems like decimal point misalignments in order amounts are corrected and recorded by the program; more complex situations, such as conflicting discount rules, generate detailed task cards for human review. For example, when a promotional activity conflicts with the payment system, the system not only highlights the conflicting clause in red but also provides relevant case references to help reviewers quickly pinpoint the root cause of the problem. This approach ensures both the efficiency of routine processing and the accuracy of critical stages through human oversight.

(2) Core challenges and solutions

1)The problem of coping with high frequency rule update

E-commerce promotion rules are frequently updated and highly complex, making it difficult for

traditional audit systems to quickly adapt to new rules. The solution is to develop a graphical rule configuration platform where business personnel can directly drag and drop components to set validation logic. Introduce a pre-check function for rule conflicts to automatically identify risk points when different promotional strategies overlap.

2) Multi-system data consistency problem

There is a delay in data synchronization for systems such as orders, payments, and risk control, and the field definition standards are not unified. The solution involves building a data traceability system to achieve cross-system data calibration through log timeline alignment and field mapping tables. An intelligent error correction module is set up to automatically trigger repair processes for abnormal data.

3) The problem of dividing the boundary between human and machine collaboration

To address the challenge of balancing automation and manual processing, the following measures are taken: risk classification management, establishing business scenario risk rating standards, and specifying handling methods for different risk levels. For example, routine high-frequency operations are automatically handled by the system, while transactions involving complex rules or large amounts are automatically referred to human review; visual decision support, developing interactive operation dashboards that intuitively display the credibility of automated processing results through color coding (green/yellow/red). For instance, abnormal orders marked in red automatically link to original data snapshots, facilitating quick verification of key fields by humans; dynamic adjustment mechanisms, continuously optimizing risk assessment thresholds based on historical review results. When the manual correction rate for a certain type of operation consistently falls below the set threshold, the system automatically expands the scope of automated processing for that scenario, forming an evolutionary closed loop of "machine learning-human calibration-rule iteration."

(3) Practical results and long-term value

This system is like equipping business processes with round-the-clock monitors, setting up automatic checkpoints at every stage from order generation to fund settlement. Whenever suspicious operations such as a single user snatching hundreds of items are detected, the transaction is immediately frozen and a verification list is generated, moving the risk interception window from end-of-month audits to the moment the transaction occurs. For instance, during rush buying hours in the early morning, the system can automatically identify abnormal order accounts, increasing the processing speed of problematic orders by five times, while fully documenting each step for easy review at any time.

The new model transforms risk control from "settling accounts after the fact" to " preparing in advance," embedding common vulnerabilities into the security fence of marketing activities. By analyzing typical issues from historical promotions, it converts error-prone discount rules into warning lines that the system can recognize. During last year's Singles' Day, when a full-reduction promotion was at risk of a flaw, the system promptly issued an alert and synchronized it with the operations team to adjust the plan, preemptively intercepting most potential problems. This shift is akin to turning a fire brigade into a fire prevention team, redirecting auditors from frantically

patching vulnerabilities to more valuable risk prediction.

6. Conclusion

Taking the "Billion Subsidy" project as a case study, this project delves into the implementation scenarios of robotic process automation in e-commerce revenue auditing. In response to the complex situation where platform promotional rules are updated more than three times a month, daily order peaks exceed 2 million, and five major business systems need to be integrated, three core modules have been developed: a dynamic rule configuration tool for rapid strategy synchronization, full-chain data tracking to ensure traceability from order placement to settlement, and an intelligent validation model for automatically identifying data conflicts across systems. This solution significantly increases audit coverage during major promotions and markedly reduces error rates.

Actual operational data shows that automated auditing has significantly reduced the time spent on reconciling single transactions. Through a drag-and-drop rule configuration interface, operations personnel can independently complete most strategy adjustments, avoiding communication losses due to technical team involvement. The system uses full-scan instead of manual sampling, identifying 12,000 abnormal orders during the "618 Mega Sale," with 43% being new arbitrage patterns. By establishing a three-tier warning mechanism—red, yellow, and blue—the risk management node is moved forward to the event phase, reducing direct losses from return fraud alone by over 8 million yuan.

The "rule configuration standardization-process execution automation-anomaly handling collaboration" methodology accumulated in the project has been compiled into a replicable implementation manual. When validated in new scenarios such as cross-border customs clearance and live-streaming sales, the dual-track mechanism of "system pre-review + manual recheck" significantly enhances audit efficiency in complex situations. A specially designed sandbox testing environment allows for simulating business changes in an isolated space, eliminating 41% of potential risks in advance. These practices provide new directions for the evolution of intelligent auditing —— The next step will be to explore using machine learning to predict wool-pulling behavior, building a more forward-looking risk control system

References

Deng, Y. G., Zhang, F., Dong, B. T. et al. (2025). Research on power grid operation data aggr egation based on RPA technology. *Northeast Electric Power Technology*, *46*(02), 34-36.

Keast, V. (2025). Beyond the random phase approximation (RPA): First principles calculation of the valence EELS spectrum for KBr including local field, quasiparticle, excitonic and spi n orbit coupling effects. *Ultramicroscopy*, *268*, 114070-114070.

Liu, C. H., & Zeng, L. Z. (2025). Construction of an Integrated Platform for Accounting Treat ment of Enterprise Data Resources Based on "LLM+RPA". *Accounting Monthly*, *46*(01), 26

-32.https://doi.org/10.19641/j.cnki.42-1290/f.2025.01.004

Mao, X., Xu, J., Jiang, J. et al. (2024). Iterative crRNA design and a PAM-free strategy enabled an ultra-specific RPA-CRISPR/Cas12a detection platform. *Communications Biology*, *7*(1), 1454-1454

Møller, S. H. C., & Sauer, A. P. S. (2025). RPA(D) and HRPA(D): Calculating NMR Spin-Spin Coupling Constants in Free Amino Acid Residues. *Magnetic resonance in chemistry*: MRC, *63*(4), 328-345.

Pang, Y., Song, J., Jiang, L. et al. (2025). Development and Evaluation of a Rapid Visualisation Detection Method for Ameson portunus Based on RPA-LFD. *Journal of fish diseases*, e14096.

Qin, Y. H., Lu, S. H., Yu, H. W. et al. (2025). Establishment and Application of a Highly Sensitive and Rapid Detection Method for Tobacco Light Green Mosaic Virus Based on RT-RPA-LFD Technology [J/OL]. *Acta Phytopathologica Sinica*, 1-10. https://doi.org/10.13926/j.cnki.apps.001365.

Success story about Robotic Process Automation (RPA) in Germany. (2025). *M2 Presswire*.

Yang, B., Yang, W., & Zha, A. Q. (2024). Research on BOM System Automation Entry Based on AI and RPA Combination. *Computer Knowledge and Technology*, *20*(36), 20-23. https://doi.org/10.14004/j.cnki.ckt.2024.1857.

Zhang, B. C. (2025). Research on the Implementation Effect of RPA in Financial Shared Service Center. *China Management Informatization*, *28*(05), 62-65.

Zhang, M. (2025). Research on High-quality Employment of Audit Talents Based on RPA Technology. *China Collective Economy*, (01), 137-140. https://doi.org/10.20187/j.cnki.cn/11-3946/f.2025.01.032