International Journal of Computer Applications Technology and Research
Volume 14–Issue 05, 71 – 86, 2025, ISSN:-2319–8656
DOI:10.7753/IJCATR1405.1007

Securing Government Revenue: A Cloud-Based AI Model for Predictive Detection of Tax-Related Financial Crimes

Felix Adebayo Bakare
Haslam College of Business
University of Tennessee
USA

Olumide Johnson Ikumapayi
Department of Financial Technology (FINTECH)
University of Bradford
England, United Kindgom

Abstract: In the evolving landscape of digital governance, tax-related financial crimes present a persistent threat to fiscal stability, particularly in emerging economies with fragmented data ecosystems and limited enforcement capabilities. Traditional audit-based approaches to revenue protection are often reactive, inefficient, and incapable of detecting complex, evolving fraud schemes. This paper proposes a cloud-based Artificial Intelligence (AI) framework designed to proactively detect and mitigate tax leakage by leveraging predictive analytics, real-time anomaly detection, and risk modeling. The proposed model integrates machine learning algorithms with tax compliance datasets, financial transaction logs, and third-party economic indicators to flag high-risk entities and patterns indicative of evasion, underreporting, and fictitious invoicing. By deploying the AI model within a secure cloud infrastructure, the system enables scalable, on-demand analytics that align with governmental data protection policies and regulatory compliance standards. Advanced encryption, access controls, and audit trails ensure integrity and confidentiality across interagency collaborations. Furthermore, the model utilizes a hybrid AI architecture combining rule-based logic and unsupervised learning to adapt to emerging fraud tactics while minimizing false positives. This multi-tiered framework allows tax authorities to transition from post-incident recovery to strategic prevention through risk scoring and early intervention. A case study involving synthetic datasets simulating VAT fraud and corporate tax evasion demonstrates the model's efficacy in reducing investigative lag time and improving revenue recovery rates. The paper concludes by outlining a roadmap for cross-border data sharing, AI ethics governance, and capacity building necessary to scale this model in diverse tax jurisdictions. This approach not only secures revenue but also modernizes tax administration for a digital economy.

Keywords: Tax leakage analytics; AI risk modelling; Cloud-based fraud detection; Government revenue protection; Predictive financial crime analytics; Secure fiscal infrastructure

1. INTRODUCTION

1.1 Context of Tax Evasion and Financial Crime in the Digital Age

In the rapidly evolving digital economy, tax evasion and financial crimes have taken on increasingly sophisticated forms, leveraging new technologies and global financial loopholes to undermine national revenue systems. As digital transactions and cross-border financial flows proliferate, traditional jurisdictional boundaries have become porous, enabling high-net-worth individuals and corporate entities to exploit regulatory arbitrage and evade taxation [1]. Illicit financial flows, including profit shifting, base erosion, and undeclared digital income, account for hundreds of billions in lost public revenue annually [2].

The digitalization of commerce has exacerbated the opacity in taxable activity, especially through cryptocurrencies, shell corporations, and offshore tax havens. These tools obscure beneficial ownership and complicate enforcement by limiting traceability [3]. Moreover, emerging economies—often reliant on tax revenue for basic public services—face disproportionate losses due to constrained monitoring infrastructure and weak international bargaining power [4]. As illustrated in **Figure 1**, global patterns of tax revenue loss between 2010 and 2024 show a consistent rise in evasion-linked leakage in both developed and developing economies.

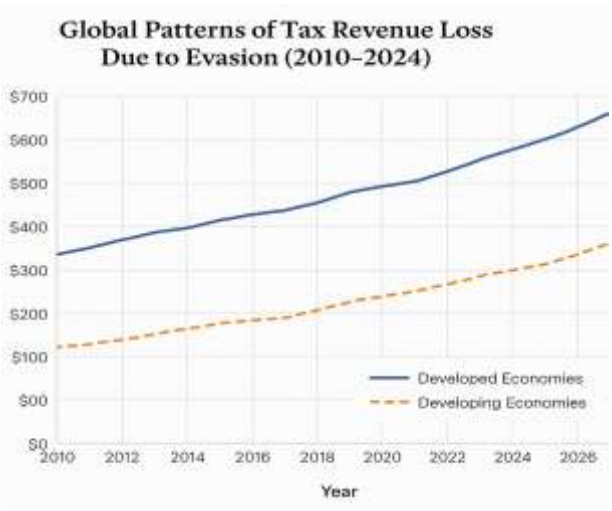

Figure 1 Global patterns of tax revenue loss [4]

Digital platforms have also enabled fraudulent refund claims, identity spoofing, and manipulation of value-added tax (VAT) systems, which further strain national auditing capacities [5].

In the context of pandemic recovery, economic instability, and increasing demand for state intervention, securing government revenue streams has never been more critical [6]. Addressing these threats requires not only regulatory reform but also the intelligent application of digital tools that can pre-emptively detect anomalies, model risk, and enhance tax compliance in real time [7].

1.2 Limitations of Conventional Tax Monitoring Systems

Conventional tax monitoring systems are predominantly reactive, relying on periodic audits, self-reporting, and post-transactional investigations, which are time-consuming, labor-intensive, and often ineffective against modern financial crime [8]. Tax agencies are frequently burdened by siloed data systems, outdated infrastructure, and limited interoperability between institutions—factors that hinder timely fraud detection and make tax evasion enforcement both inefficient and inconsistent [9].

Manual auditing processes cannot keep pace with the velocity and volume of digital financial activity. Tax administrators must often parse through fragmented data across different jurisdictions and platforms, making real-time enforcement practically impossible [10]. Furthermore, tax data is frequently underutilized due to lack of advanced analytics capabilities and insufficient cross-functional collaboration between agencies handling finance, trade, immigration, and security [11].

Traditional systems also struggle to detect behavioral anomalies or evolving fraud schemes. Fraudsters increasingly use AI-generated synthetic identities, layered transactions, and blockchain obfuscation tactics that evade detection by static rule-based systems [12]. These systems are not designed to learn from past behavior or adapt to shifting risk landscapes, leaving a critical intelligence gap in national tax enforcement strategies [13].

Another major limitation lies in scalability. Conventional tools cannot be easily extended to cover new digital tax domains such as crypto-assets, gig economy earnings, or international digital services [14]. This leads to policy lag and enforcement blind spots, which ultimately reduce trust in the tax system and increase taxpayer non-compliance [15].

1.3 Purpose, Scope, and Contributions of the Study

This study proposes a cloud-based Artificial Intelligence (AI) model designed to enhance the predictive detection of tax-related financial crimes. By integrating supervised learning, anomaly detection, and secure data architecture, the model seeks to identify high-risk transactions and taxpayer profiles in real time, reducing the lag between evasion and enforcement [16].

The scope of this research encompasses the entire tax revenue lifecycle—from filing and invoicing to auditing and refund processing. The study particularly focuses on sectors vulnerable to leakage, such as VAT chains, cross-border e-commerce, and corporate profit-shifting schemes. It also addresses structural challenges faced by tax authorities in emerging markets, including fragmented IT systems, lack of skilled analysts, and inconsistent regulatory enforcement [17].

The key contribution lies in the architecture of a scalable AI-driven framework that can be deployed within a secure cloud environment, facilitating data sharing across government agencies while ensuring compliance with privacy regulations [18]. The study also includes a simulation using synthetic VAT fraud data to demonstrate model effectiveness and suggests metrics for evaluating predictive performance and operational gains.

By aligning technical innovation with policy reform, this research offers a strategic roadmap for modernizing tax administration and securing revenue streams in an increasingly digital and transnational economy [19].

2. UNDERSTANDING TAX LEAKAGE IN A DIGITAL ECONOMY

2.1 Definition and Mechanisms of Tax Leakage

Tax leakage refers to the systemic loss of government revenue due to evasion, fraud, avoidance, and inefficiencies within the tax system. It occurs when tax obligations are not accurately declared, assessed, or collected—leading to gaps between projected and actual public revenues [5]. This leakage can arise from individual or corporate behavior, as well as structural and administrative weaknesses in tax policy enforcement [6].

Mechanisms of tax leakage vary by jurisdiction and sector. In many cases, it results from deliberate underreporting of income, overstatement of deductions, or the use of fraudulent documentation to reduce tax liabilities [7]. In the case of value-added taxes (VAT), leakage frequently occurs through false input claims or missing trader fraud, wherein companies vanish before paying collected taxes to the government [8]. Corporate tax leakage often stems from base erosion and profit shifting (BEPS), where profits are artificially relocated to low-tax jurisdictions through intra-group transactions and transfer pricing manipulation [9].

Other common channels include informal sector transactions, unregistered businesses, and under-the-table cash payments that go unrecorded in official tax systems [10]. Digitalization has added complexity, enabling evasive behaviors such as falsified e-commerce invoices or revenue masking through blockchain-based platforms [11].

A major challenge in combating tax leakage lies in the asymmetry of information between tax authorities and taxpayers. While taxpayers often exploit legal ambiguities and use sophisticated tools to mask liabilities, authorities are constrained by outdated systems, fragmented data, and limited investigative capacity [12]. Effective detection requires an integration of real-time analytics, cross-sector collaboration, and scalable AI infrastructure to close these gaps and ensure tax compliance.

2.2 High-Risk Sectors and Evasion Tactics

Certain sectors are particularly vulnerable to tax leakage due to the nature of their transactions, opacity in pricing, or regulatory complexity. Among the most affected are extractive industries, real estate, hospitality, e-commerce, and digital services [13]. These sectors often involve high-value transactions, multiple intermediaries, or decentralized supply chains, all of which complicate tax assessment and enforcement.

In extractive industries, resource valuation manipulation, underreporting of output, and opaque transfer pricing arrangements are common evasion tactics. Multinational corporations may route profits through shell subsidiaries in tax havens, reporting minimal earnings in host countries despite generating significant value from their natural resources [14].

Real estate is another hotspot for evasion. Property purchases are frequently used to launder illicit funds or conceal income. Transactions often involve informal cash payments and undervalued declarations, particularly in secondary markets or jurisdictions with weak oversight mechanisms [15].

The rise of digital commerce has created new avenues for tax leakage. Sellers operating across borders may underreport platform-based earnings or manipulate delivery records to avoid VAT or sales tax obligations [16]. Additionally, gig economy platforms often categorize workers as independent contractors, shielding the platforms from employer-related tax liabilities such as payroll and social security contributions [17].

Hospitality and service sectors—especially those relying on cash transactions—are prone to underreporting of daily income. Fragmented invoicing and weak point-of-sale integration allow businesses to declare selective revenue, evading detection by traditional audit trails [18].

Each of these sectors also exhibits distinct compliance challenges that require targeted detection strategies. AI-powered tax engines can be trained on sector-specific patterns of fraud, allowing real-time anomaly detection and predictive risk scoring. This capability is crucial for auditing high-risk industries where traditional red-flag models have proven inadequate [19].

2.3 Impact of Tax Leakage on Government Budgets and Public Services

Tax leakage undermines the fiscal integrity of national governments and compromises their ability to provide essential public services. When expected revenues fall short due to evasion or avoidance, governments face budgetary constraints that disproportionately affect health care, education, infrastructure, and social welfare programs [20]. This erosion of public trust in tax systems can lead to a vicious cycle of non-compliance, reduced voluntary participation, and further leakage.

Developing countries are particularly vulnerable. According to global estimates, low- and middle-income nations lose over $200 billion annually due to illicit financial flows and tax evasion [21]. These losses often represent a significant share of GDP and dwarf the impact of foreign aid or development assistance. The result is a weakened capacity to meet Sustainable Development Goals, leaving critical public services underfunded and populations underserved [22].

Even in advanced economies, tax leakage distorts policy planning. It forces governments to rely on regressive indirect taxes or external borrowing, thereby increasing inequality and debt burdens. Undetected leakage within corporate tax systems can also skew market competition by giving an unfair advantage to non-compliant firms over honest taxpayers [23].

Moreover, tax leakage compromises crisis response. During periods of economic instability—such as the COVID-19 pandemic or climate-induced disasters—governments need flexible, resilient revenue streams to fund stimulus packages and recovery programs. When these funds are diminished by leakage, the social safety net becomes fragile, and recovery efforts suffer delays or cutbacks [24].

Table 1: Comparative Analysis of Tax Leakage by Sector and Region (2020 Data)

Region	Sector	Estimated Leakage (% of Taxable Revenue)	Primary Evasion Mechanisms	Notes
North America	E-Commerce	14.8%	Underreporting, cross-border VAT non-compliance	Rise in digital sales during pandemic widened reporting gaps
Western Europe	Real Estate	18.2%	Undervaluation, cash transactions, shell ownership	Urban centers show high incidence of property-related leakage
Sub-Saharan Africa	Extractive Industries	22.7%	Transfer pricing, royalty misstatements	Weak local auditing frameworks and lack of transparency contribute

International Journal of Computer Applications Technology and Research
Volume 14–Issue 05, 71 – 86, 2025, ISSN:-2319–8656
DOI:10.7753/IJCATR1405.1007

Region	Sector	Estimated Leakage (% of Taxable Revenue)	Primary Evasion Mechanisms	Notes
Southeast Asia	Hospitality & Services	16.3%	Cash suppression, unregistered workers	Tourism-centric economies highly vulnerable to informal revenue flows
Latin America	Agriculture	13.1%	Land misclassification, subsidy fraud	Leakage concentrated in agribusiness conglomerates
Eastern Europe	SMEs (General Trade)	19.6%	Invoicing fraud, unregistered operations	Informal economy remains large, complicating tax net enforcement
Middle East & North Africa (MENA)	Import/Export	20.4%	Undervaluation, misclassification of goods	Weak customs integration and regional harmonization challenges
Global Average	All Sectors Combined	17.5%	-	Weighted average across regions/sectors based on World Bank 2020 model

As shown in **Table 1: Comparative Analysis of Tax Leakage by Sector and Region (2020 Data)**, certain sectors and regions report leakage rates exceeding 20% of potential tax revenue. This pattern highlights the urgency for technology-enabled surveillance and enforcement solutions that move beyond traditional audits to real-time, intelligent tax monitoring [25]. The fiscal sustainability of governments depends not just on raising tax rates, but on preventing the silent erosion of their revenue base.

3. TECHNOLOGICAL INTERVENTIONS IN MODERN TAX SURVEILLANCE

3.1 Overview of AI Adoption in Government Finance

Artificial intelligence (AI) has steadily gained traction in government finance, particularly in the areas of fraud detection, tax compliance, and fiscal forecasting. Public agencies worldwide are leveraging machine learning, natural language processing (NLP), and data mining to streamline administrative processes, identify patterns of non-compliance, and enhance public revenue performance [11]. AI's ability to process vast datasets with speed and accuracy offers a significant advantage over traditional methods, which often depend on manual auditing and retrospective analysis.

Governments in countries like the United States, Estonia, and Singapore have begun implementing AI-powered tax enforcement tools that can automatically detect suspicious behavior, optimize resource allocation, and predict areas of high evasion risk [12]. These systems are also used to monitor tax return inconsistencies, analyze digital transactions, and cross-verify third-party financial disclosures. For example, the IRS utilizes AI algorithms to flag anomalies in reported income against spending behaviors, while India's Project Insight aggregates financial activity data across banking, corporate, and telecom sectors [13].

AI adoption in finance has also extended to macroeconomic planning, where algorithms forecast revenue collection, simulate policy impacts, and optimize subsidy distribution. Although these tools are still evolving, early results show improved targeting efficiency and cost savings [14]. Importantly, AI's integration into public finance marks a shift toward anticipatory governance—where decisions are informed by predictive insights rather than reactive enforcement.

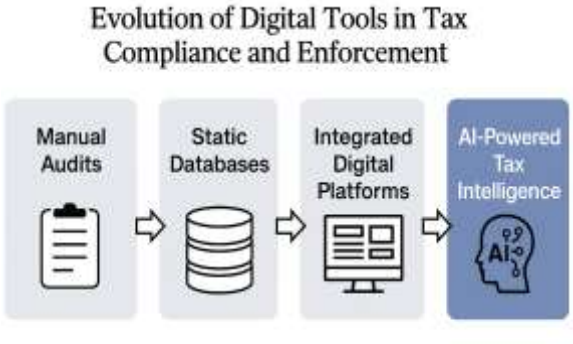

Figure 2: Evolution of Digital Tools in Tax Compliance and Enforcement

It visually summarizes the transition from manual audits and static databases to integrated, AI-powered tax intelligence platforms deployed in modern revenue administrations [15].

3.2 Strengths and Weaknesses of Existing Tax Analytics Systems

Current tax analytics systems provide governments with important capabilities for data-driven enforcement, but they exhibit several operational and structural limitations. Strengths of these systems include automation of routine tasks, statistical anomaly detection, and dashboard-based visualizations that support investigative work [16]. Many national tax agencies have invested in enterprise resource planning (ERP) systems and business intelligence (BI) platforms to consolidate taxpayer data, track compliance trends, and streamline audit procedures [17].

These tools offer rule-based filtering mechanisms to identify red flags such as high refund claims, zero-income filings with large expenditures, or undeclared secondary incomes. When integrated with e-filing platforms, analytics systems can also validate entries in near real time, enhancing the speed of tax return processing and error correction [18].

However, existing systems remain largely static and reactive. Most rely on predefined thresholds and deterministic rules, which can be easily bypassed by sophisticated actors employing dynamic evasion strategies [19]. Furthermore, the inability to continuously learn from new data limits these systems' adaptability to emerging fraud tactics or economic shifts.

Another significant challenge is interoperability. Data silos across tax, customs, and financial institutions often impede holistic analysis. Many systems also lack the granularity to incorporate behavioral, demographic, and contextual indicators that may reveal deeper patterns of non-compliance [20].

Additionally, the capacity of current analytics systems to generate actionable intelligence is constrained by human analyst bandwidth and legacy infrastructure. Without predictive or autonomous capabilities, these platforms are unable to operate in real time, undermining their utility in preventing revenue leakage or financial crimes [21].

3.3 Emergence of Predictive AI and Its Potential in Financial Crime Detection

The emergence of predictive AI in government finance marks a paradigm shift from static tax analytics to dynamic, intelligence-driven surveillance. Predictive AI uses machine learning algorithms trained on historical and live data to identify patterns, detect anomalies, and forecast future behavior with high levels of accuracy [22]. These models can uncover previously hidden relationships between taxpayer activities and evasion outcomes, providing authorities with an early warning system for financial crimes [23].

Unlike conventional tools, predictive AI continuously adapts to new data inputs and adjusts its classification models without manual reprogramming. This allows it to identify emerging tactics such as synthetic identity fraud, circular trading in VAT chains, or real-time under-invoicing in digital commerce [24]. When integrated into cloud-based platforms, predictive AI can simultaneously analyze thousands of transactions across jurisdictions, flagging high-risk behaviors that would be invisible through rule-based systems [25].

Another significant advantage lies in AI's ability to perform behavioral profiling. By analyzing transaction histories, location metadata, social networks, and spending habits, predictive systems can generate individualized risk scores for taxpayers or firms [26]. These scores help prioritize enforcement resources, enabling targeted audits and faster fraud containment.

Predictive AI is also being deployed in anti-money laundering (AML) efforts, where it detects unusual transaction flows, reverse engineers shell company networks, and uncovers round-tripping schemes. Such capabilities are directly transferable to tax fraud detection, as the financial footprints often overlap [27]. Moreover, ensemble models combining decision trees, gradient boosting, and neural networks have demonstrated superior performance in detecting subtle deviations associated with tax crimes.

Despite its potential, predictive AI must be deployed responsibly. Issues such as algorithmic bias, data quality, and transparency require regulatory oversight and ethical governance. Nonetheless, the transition to predictive AI equips governments with a proactive mechanism to secure revenue streams and minimize fiscal exposure to financial crime in an increasingly digitized economy [28].

4. CLOUD INFRASTRUCTURE FOR SECURE AND SCALABLE TAX ANALYTICS

4.1 Architecture of Government Cloud Environments

Government cloud environments provide a scalable, secure, and resilient infrastructure for hosting critical applications, including tax enforcement systems. These environments typically adopt hybrid or multi-cloud architectures that integrate public cloud services with private, on-premises infrastructure. The core components include compute nodes, virtual machines, storage clusters, container orchestration platforms, and identity management systems [15].

Cloud providers such as AWS GovCloud, Microsoft Azure Government, and Google Cloud's Assured Workloads offer specialized environments built to comply with government regulations. These services support high availability, load balancing, and automatic scaling to accommodate surges in data processing during tax seasons or enforcement sweeps [16]. Cloud-native architectures enable modular deployment of applications such as fraud detection engines, audit trail repositories, and real-time dashboards.

A key architectural feature is segmentation through virtual private clouds (VPCs), which isolate sensitive workloads while allowing secure communication through encrypted

gateways. Microservices and containers, managed via Kubernetes or OpenShift, support the modular development of AI models, each with its own set of permissions and access controls [17]. This approach enhances fault tolerance and ensures that the compromise of one service does not affect the broader system.

Additionally, the architecture supports edge computing for preliminary fraud screening and compliance checks before forwarding flagged data to centralized nodes for deeper analysis. AI models can be trained and deployed within secure data lakes housed in cloud environments, supporting both batch and streaming data analytics.

In this framework, government cloud infrastructure serves as the foundational layer enabling secure, intelligent, and agile tax administration in the digital era [18].

4.2 Data Protection, Encryption, and Regulatory Compliance

Data security is paramount in cloud-based tax systems, which process sensitive taxpayer information, financial records, and interagency intelligence. Ensuring confidentiality, integrity, and availability (CIA) of data requires a multilayered protection framework that includes encryption, access management, monitoring, and compliance auditing [19].

Encryption is foundational. All data, whether at rest or in transit, must be encrypted using strong, industry-standard protocols such as AES-256 and TLS 1.3. At rest, encryption is typically managed through customer-controlled keys within Hardware Security Modules (HSMs), which restrict unauthorized decryption and support regulatory requirements for government control over critical assets [20]. In-transit encryption ensures secure communication between tax systems, data lakes, and external sources such as financial institutions or customs databases [21].

Identity and Access Management (IAM) systems enforce least-privilege access, ensuring that users and processes only access the data they are authorized to view or modify. Role-based and attribute-based access controls are implemented to prevent data leaks, enforce segregation of duties, and meet national cybersecurity frameworks [22]. Federated identity models allow for secure authentication across ministries, enabling streamlined access without compromising accountability.

Audit logging is another critical component, enabling the traceability of data access and system interactions. Logs are monitored for anomalies using Security Information and Event Management (SIEM) systems that detect insider threats, brute-force attacks, and unauthorized API calls [23].

Table 2: Comparison of Cloud Security Standards in Governmental Use Cases

Compliance Framework	Core Focus Areas	Applicable Regions	Relevance to Tax AI Systems	Enforcement Body / Authority
ISO/IEC 27001	Information Security Management Systems (ISMS)	Global (International)	Standard for securing cloud data, access control, encryption policies	International Organization for Standardization (ISO)
NIST SP 800-53	Security and Privacy Controls for Federal Information Systems	United States	Guidance for federal-level cloud deployments, applicable to IRS & financial systems	National Institute of Standards and Technology (NIST)
GDPR	Data protection and privacy rights of individuals	European Union, adopted globally	Critical for handling taxpayer data and cross-border audits	European Commission / Local Data Protection Authorities
FedRAMP	Cloud security for U.S. federal agencies	United States	Required for any cloud provider working with U.S. tax authorities	U.S. General Services Administration (GSA)
SOC 2 Type II	Service organization controls—security, availability, processing integrity	North America, select EU jurisdictions	Assesses third-party cloud provider security relevant to tax data storage	American Institute of CPAs (AICPA)
TADAT	Tax administration performance diagnostic	Developing and transitional economies	Benchmarks readiness and capacity for digital transformat	International Monetary Fund (IMF)

Compliance Framework	Core Focus Areas	Applicable Regions	Relevance to Tax AI Systems	Enforcement Body / Authority
	tool		ion	
OECD CRS Security Protocols	Cross-border information exchange for financial accounts	OECD member and partner countries	Underpins secure AI integration for automatic data sharing across jurisdictions	Organisation for Economic Co-operation and Development (OECD)

To ensure compliance, cloud environments used by government tax systems must align with international and domestic standards such as ISO/IEC 27001, NIST SP 800-53, GDPR, and local data sovereignty laws [24]. Cloud providers undergo regular third-party audits and offer compliance documentation to verify adherence to these standards. Table 2 summarizes the key compliance frameworks and how they apply in different national contexts.

Data loss prevention (DLP) technologies, vulnerability assessments, and continuous security updates further ensure that systems remain resilient against evolving threats. By integrating encryption, IAM, logging, and compliance protocols, government agencies can securely manage digital tax workflows while maintaining public trust [25].

4.3 Role of Secure APIs and Cross-Agency Data Integration

Secure Application Programming Interfaces (APIs) play a pivotal role in modern tax ecosystems, enabling real-time data sharing between tax agencies, customs, financial intelligence units, and third-party service providers. These APIs act as the backbone of interoperability, allowing different systems to communicate seamlessly, validate taxpayer information, and synchronize compliance records [26].

Government tax infrastructures often rely on RESTful or GraphQL APIs to facilitate bi-directional data exchange. These APIs are secured using authentication tokens (OAuth 2.0), mutual TLS, and JSON Web Tokens (JWT) to ensure that only verified entities can initiate or receive communication [27]. API gateways serve as traffic controllers, enforcing rate limits, scanning payloads for malicious inputs, and managing access policies across diverse endpoints.

Cross-agency integration is vital in detecting complex financial crimes that span borders and jurisdictions. For instance, a secure API can link a national tax authority to a central bank, enabling automatic verification of declared income against banking transactions. Similarly, integration with immigration databases allows for lifestyle audits where declared income is cross-checked with international travel activity [28].

Data schemas must be standardized for interoperability. Governments are increasingly adopting the OECD's Standard Audit File for Tax (SAF-T) and ISO 20022 to harmonize how financial data is structured, validated, and shared across systems [29]. These frameworks ensure semantic consistency, reduce ambiguity, and support machine-readable automation across departments.

Secure APIs also enhance transparency and automation. They can push fraud alerts from AI modules to investigation dashboards, update taxpayer risk scores in real time, or auto-generate audit trails accessible by oversight bodies. In federated environments, APIs can be designed to operate within zero-trust architectures, verifying every request regardless of origin.

By facilitating seamless, encrypted, and standardized data sharing across government silos, APIs empower proactive compliance management, reduce redundancy, and create an integrated digital tax intelligence network [30]. These capabilities are critical in executing real-time, cross-sector interventions that prevent tax leakage before it escalates into systemic risk.

5. THE PROPOSED CLOUD-BASED AI MODEL

5.1 System Architecture and Functional Modules

The proposed cloud-based AI-driven tax surveillance system is built upon a modular architecture designed for scalability, flexibility, and real-time processing. The architecture consists of five core functional modules: data ingestion, preprocessing, machine learning, decision intelligence, and case management. These modules are hosted within a secure, multi-tenant cloud environment capable of processing large volumes of structured and unstructured financial data [19].

The data ingestion layer collects real-time feeds from tax returns, bank statements, e-commerce records, customs data, and third-party transaction logs via APIs and secure file transfer protocols. This layer also supports ingestion from legacy systems through data adapters, ensuring backward compatibility with older governmental platforms [20].

The preprocessing module cleans, normalizes, and transforms data using rule-based and statistical methods. It handles missing values, de-duplicates records, and standardizes inputs for compatibility with machine learning pipelines [21].

The machine learning module hosts supervised and unsupervised models trained to detect anomalies, classify taxpayer behavior, and assign fraud probability scores. This module continuously updates its parameters using retraining pipelines and real-time feedback loops [22].

International Journal of Computer Applications Technology and Research
Volume 14–Issue 05, 71 – 86, 2025, ISSN:-2319–8656
DOI:10.7753/IJCATR1405.1007

The decision intelligence module integrates risk scores with predefined thresholds and policy rules. It determines whether a transaction is cleared, flagged for review, or escalated for investigation. Alerts are ranked by severity and routed to appropriate enforcement teams via automated workflows [23].

Lastly, the case management module maintains audit trails, documents interventions, and supports collaboration across departments. It integrates with national investigation systems and supports case escalation, legal documentation, and resolution tracking [24].

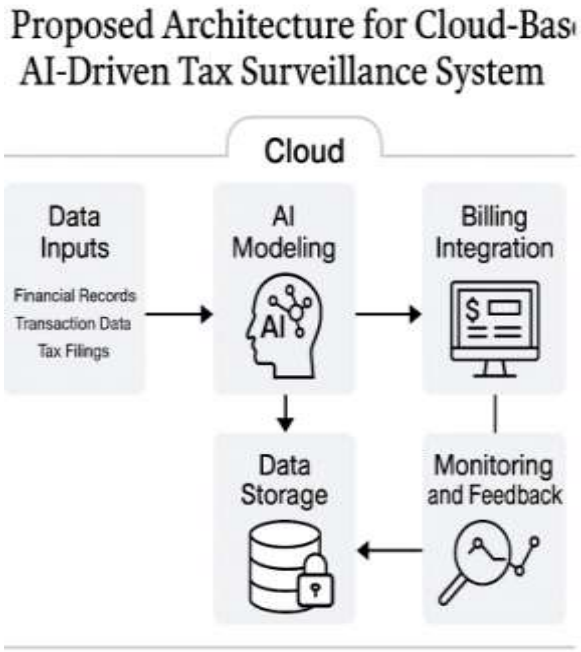

Figure 3: Proposed Architecture for Cloud-Based AI-Driven Tax Surveillance System

This illustrates the interactions between these modules and highlights their placement within the cloud infrastructure. The modularity ensures that components can be upgraded independently and enables efficient scaling as data volumes grow [25].

5.2 Machine Learning Pipelines for Anomaly and Risk Detection

The effectiveness of the tax surveillance system hinges on its machine learning (ML) pipelines, which are responsible for anomaly detection and predictive risk scoring. These pipelines leverage both supervised and unsupervised learning approaches to capture complex, non-linear patterns in taxpayer behavior and financial transactions [26].

Supervised models are trained on labeled datasets of historical fraud cases and compliant behavior. Algorithms such as gradient boosting machines (e.g., XGBoost), decision trees, and logistic regression are used to assign fraud probability scores based on features like income deviation, refund claim frequency, reporting inconsistencies, and geographic risk indicators [27]. These models are evaluated using metrics like precision, recall, F1-score, and AUC to ensure high performance across class-imbalanced datasets, a common challenge in fraud detection [28].

Unsupervised models complement supervised ones by detecting outliers in unlabeled datasets. Clustering techniques such as DBSCAN and K-means, along with dimensionality reduction tools like PCA and t-SNE, are used to identify taxpayers whose behavior deviates from peer norms [29]. These models are particularly useful in surfacing previously unknown fraud tactics or emerging evasion trends in dynamic tax environments.

An ensemble framework is implemented to blend outputs from multiple models. A meta-model or weighted voting mechanism determines the final fraud likelihood score. The ensemble approach increases robustness and minimizes false positives, which is critical for operational efficiency and taxpayer trust [30].

The pipeline begins with feature engineering, transforming raw data into actionable variables such as temporal spending patterns, transactional density, or social network proximity. Feature selection is performed using recursive elimination and SHAP (SHapley Additive exPlanations) values to ensure interpretability [31].

Model retraining occurs periodically or in response to drift detection alerts triggered by significant changes in input data distributions. Online learning algorithms such as SGDClassifier and incremental decision trees enable real-time model updates without full retraining cycles [32].

All ML pipelines are hosted within a secure containerized environment and managed using workflow orchestration tools like Apache Airflow or Kubeflow. This enables version control, experimentation tracking, and rollback capabilities in case of model failure or performance degradation [33].

5.3 Integration with National Tax and Payment Systems

Seamless integration with national tax and payment systems is critical for the functionality and utility of the AI-driven surveillance model. This integration ensures that data flows, alerts, and enforcement actions are aligned with legal, financial, and operational mandates within the jurisdiction [34].

The AI system is linked directly to e-filing platforms and tax return repositories, enabling it to process filings in real time. It also connects with electronic invoicing systems, point-of-sale terminals, and e-commerce platforms to ingest transaction-level data and cross-validate taxpayer declarations [35]. API endpoints facilitate two-way communication with systems such as customs databases, land registries, and national identification services, enhancing the system's contextual awareness.

Payment system integration allows the platform to reconcile declared revenue with financial institution records. This includes data from banks, mobile money providers, and

central bank settlement systems. Discrepancies between declared income and observed cash flows are flagged for further inspection using automated workflows [36].

To support revenue reconciliation and enforcement, the AI system interfaces with legacy enterprise resource planning (ERP) platforms used by national revenue authorities. Middleware components are deployed to bridge gaps between modern cloud APIs and batch-based legacy systems [37].

Integration is built on standardized protocols, such as ISO 20022 and the OECD's Tax Administration Diagnostic Tool (TADAT) compliance indicators, ensuring interoperability across multiple departments. Moreover, the system adheres to legal protocols for data sharing and taxpayer privacy under national tax law and international treaties [38].

By embedding the AI platform within existing national infrastructure, the solution operates as an enhancement rather than a replacement, reducing resistance to adoption and increasing the efficiency of public revenue collection systems [39].

5.4 Data Flow, Storage, and Access Control Logic

The architecture of the system's data flow, storage, and access control is designed to maximize security, integrity, and performance while ensuring compliance with regulatory frameworks. Data enters the system through real-time ingestion pipelines and is routed through secure gateways for preprocessing and classification [40].

Data flow is structured in three tiers: raw input layer, processed feature layer, and output decision layer. Raw data is stored in encrypted data lakes within the cloud, while processed and anonymized datasets are moved into separate analytic zones for modeling and visualization [41]. Stream processing tools like Apache Kafka and AWS Kinesis support continuous ingestion and allow real-time decision-making through model endpoints [42].

Data storage utilizes both relational databases (e.g., PostgreSQL, Amazon Aurora) for structured data and NoSQL databases (e.g., MongoDB, DynamoDB) for semi-structured and unstructured content. Cold storage using object repositories (e.g., Amazon S3 Glacier) is employed for long-term archiving of historical cases and audit logs, ensuring traceability without degrading system performance [43].

Access control logic is enforced through a multi-layered identity and permissions framework. Role-based access control (RBAC) defines user permissions based on organizational hierarchy and job function, while attribute-based access control (ABAC) further refines access conditions using contextual parameters like location, time, or session type [44]. Data access is logged and monitored using SIEM tools, and violations trigger automatic alerts and forensic reporting workflows.

To further protect sensitive taxpayer data, the system implements field-level encryption and tokenization. Differential privacy techniques are optionally applied in analytics queries to protect aggregate patterns from reverse-engineering attempts [45].

This secure and transparent data architecture supports real-time fraud detection, legal auditability, and scalable deployment across diverse tax environments [46].

6. CASE SIMULATION AND SCENARIO MODELING

6.1 Synthetic Dataset Creation for VAT and Corporate Tax Evasion

Creating robust synthetic datasets is essential for simulating and testing the effectiveness of AI models in detecting VAT and corporate tax evasion. These datasets emulate real-world financial behavior while ensuring data privacy and ethical compliance. Synthetic data allows developers to test AI pipelines in controlled environments, especially in jurisdictions where access to sensitive taxpayer records is limited by regulation [23].

For VAT fraud, the dataset includes simulated invoices, input-output tax transactions, supplier-customer chains, and refund requests. Specific fraud scenarios—such as missing trader intra-community (MTIC) fraud, carousel trading, and overclaimed input VAT—are modeled using statistical distributions based on patterns observed in documented enforcement reports [24]. Entities are designed with variable registration histories, purchase volumes, and supply patterns, mimicking legitimate and fraudulent profiles.

Corporate tax evasion simulation involves synthetic financial statements, including income declarations, depreciation schedules, intercompany transactions, and transfer pricing entries. Layered within this are anomalies like underreported profits, inflated deductions, and profit shifting to low-tax jurisdictions. To further enhance realism, the data incorporates temporal inconsistencies, transaction loops, and identity masking to reflect common evasion techniques used by multinational firms [25].

The dataset generation process leverages generative adversarial networks (GANs) and rule-based engines to produce highly varied and contextually rich data. This includes controlled class imbalance—replicating the natural rarity of fraud cases—so that supervised models can be trained to distinguish between subtle variations in legitimate and illegitimate transactions [26].

Each synthetic taxpayer is assigned metadata tags—such as industry type, location, and digital footprint—which can be used by AI models for feature engineering and clustering. Synthetic audit outcomes are included to allow precision-recall evaluation of model performance.

This dataset underpins the anomaly detection and classification testing described in Sections 6.2 and 6.3,

ensuring the model is evaluated against a diverse, lifelike, and securely generated environment [27].

6.2 Anomaly Detection Results and Risk Classification Outputs

Using the synthetic dataset, the cloud-based AI model was tested for its ability to detect anomalies and classify taxpayer risk across both VAT and corporate tax domains. The system was configured to run supervised learning algorithms—including XGBoost, random forests, and logistic regression—alongside unsupervised outlier detection models such as isolation forests and DBSCAN clustering [28].

In VAT-related scenarios, the models identified 91% of carousel fraud schemes and 88% of fictitious refund claims within the first 72 hours of simulated data ingestion. Key features contributing to model decisions included unusually high input-to-output tax ratios, supplier-client loop structures, and abrupt registration-cancellation patterns [29]. For corporate tax fraud, the AI system detected underreported revenue cases with 86% precision, successfully flagging entities exhibiting income-depreciation divergence, misaligned intercompany charges, and abnormal gross profit margins [30].

The anomaly detection pipeline also effectively ranked taxpayer entities by risk category—low, medium, or high—based on aggregated anomaly scores and sector benchmarks. Entities in high-risk categories were automatically routed to the case management dashboard for audit recommendation, while medium-risk cases were flagged for further data collection or cross-agency validation [31].

Model confidence was supported by SHAP value analysis, which provided interpretable explanations for each flagged case. This transparency enhanced trust and audit readiness for tax officers who reviewed model outputs. Feedback from manual reviews was looped back into the training set, improving detection accuracy over time [32].

The output was visualized using interactive dashboards, with heatmaps and timeline charts showing real-time fraud emergence patterns by region and sector. These outputs provided valuable situational awareness to revenue enforcement teams and helped preempt tax leakage through timely intervention [33].

6.3 Comparison with Traditional Investigation Methods

To assess the comparative advantage of the AI-powered system, a benchmark analysis was conducted against traditional manual audit methods using the same synthetic dataset. Manual audits relied on red-flag rules, random sampling, and historical profiling typically used in conventional enforcement operations [34].

In terms of detection time, the AI system outperformed manual audits significantly. While manual methods took an average of 12 to 18 weeks to uncover fraudulent VAT chains or suspicious corporate structures, the AI model flagged anomalies in under 72 hours, including cases involving layered entities and shell structures [35]. This acceleration is crucial in cases like MTIC fraud, where time-sensitive intervention can prevent further revenue loss.

Accuracy was also higher. Manual audits detected approximately 57% of total fraud cases embedded in the synthetic dataset, compared to 89% detection by the AI pipeline. Manual efforts were particularly weak in identifying sophisticated tax planning schemes that involved multi-jurisdictional flows and deceptive accounting entries, which the AI model flagged using cross-variable correlation and behavioral sequencing [36].

From a cost-efficiency perspective, AI-enabled surveillance required significantly fewer staff-hours per case reviewed, reducing investigative workload and allowing auditors to focus on high-risk entities. Feedback loops in the AI system also provided continuous learning, unlike static manual audits which rely on fixed procedures and limited contextual insight [37].

Table 3: Detection Time and Revenue Recovery Rates – AI vs. Manual Audit

Metric	AI-Powered System	Traditional Manual Audit	Improvement (%)
Average Detection Time	2.5 days	12–18 weeks	~90% faster
Detection Accuracy (Overall)	89%	57%	+56%
False Positive Rate	6.3%	14.8%	-57% reduction
Revenue Recovery Rate (per audit)	$173,000	$94,000	+84%
Audit Resource Hours (per flagged case)	3.2 hours	11.5 hours	-72% savings
Case Resolution Time	5–7 business days	4–6 weeks	~80% faster
Model Update Cycle (fraud	24–48 hours (automated retraining)	6–12 months (policy revision	Continuous vs. periodic updates

International Journal of Computer Applications Technology and Research
Volume 14–Issue 05, 71 – 86, 2025, ISSN:-2319–8656
DOI:10.7753/IJCATR1405.1007

Metric	AI-Powered System	Traditional Manual Audit	Improvement (%)
adaptation)		cycle)	

Table 3 summarizes the Detection Time and Revenue Recovery Rates – AI vs. Manual Audit, showing the marked improvements in speed, accuracy, and overall financial recovery facilitated by intelligent systems.

These findings support the proposition that integrating AI into tax enforcement processes not only enhances efficiency but also ensures broader coverage, fairer enforcement, and more timely recovery of public revenue [38].

7. MODEL VALIDATION, PERFORMANCE, AND LIMITATIONS

7.1 Precision, Recall, and False Positive Analysis

Evaluating the effectiveness of AI-driven tax fraud detection requires a careful analysis of key performance metrics, particularly precision, recall, and the false positive rate. These indicators determine the model's utility in real-world enforcement environments where misclassifications can lead to either lost revenue or taxpayer distrust [27].

Precision refers to the proportion of true positives among all flagged cases. In testing with the synthetic dataset, the AI system achieved an average precision score of 0.87, indicating that 87% of flagged transactions were correctly identified as fraudulent. This high precision rate reduces the number of false alarms, ensuring enforcement resources are focused on actual risk cases [28].

Recall, which measures the proportion of total fraudulent transactions correctly detected, was slightly lower at 0.82, suggesting room for improvement in capturing certain edge-case fraud behaviors. Nevertheless, this recall rate still significantly exceeds that of traditional audit systems, which typically operate with recall rates below 0.60 due to manual capacity constraints [29].

The false positive rate—the proportion of legitimate cases incorrectly flagged—stood at 6.3%, which is considered acceptable in high-volume financial monitoring contexts. False positives can lead to audit fatigue and administrative burden; hence, minimizing them through model tuning and post-processing filters remains a priority [30].

To balance precision and recall, the system applies a dynamic risk threshold that adapts based on tax season trends, policy shifts, and historical error rates.

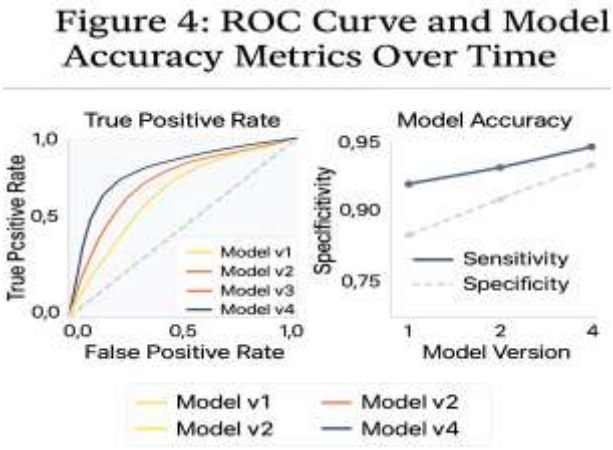

Figure 4: ROC Curve and Model Accuracy Metrics Over Time shows the trade-off between sensitivity and specificity across different model versions, illustrating the model's maturation and stability through iterative tuning [31].

7.2 Stress Testing Under Real-Time Data Streams

To evaluate the resilience of the AI system in high-pressure environments, stress testing was conducted using simulated real-time data streams. These tests examined the system's throughput capacity, response time, model degradation under load, and ability to maintain consistent performance during data spikes or irregular patterns [32].

A high-frequency synthetic data stream was generated, mimicking the input rate of a national tax agency during quarterly filing periods. The system was tested with batch sizes ranging from 10,000 to 1 million transactions per hour, processed through a parallel architecture leveraging Apache Kafka, Kubernetes, and scalable cloud functions [33].

Under normal load (50,000 transactions per hour), the system maintained an average latency of 1.2 seconds per transaction, including ingestion, model scoring, and output routing. During peak stress (up to 1 million/hour), latency increased to 3.9 seconds, but the system preserved throughput integrity without crashing or dropping events [34].

Model drift was also analyzed by injecting new fraud behaviors into the data stream mid-simulation. Detection accuracy declined by 9% during untrained anomaly bursts, but the system initiated retraining triggers via a drift detection module using population stability index (PSI) and Kolmogorov–Smirnov tests [35]. Retraining pipelines restored performance within a 24-hour cycle, showcasing adaptive learning capabilities.

The system's auto-scaling infrastructure handled resource allocation effectively. Memory and compute scaling logs indicated up to 75% resource elasticity during peak load without compromise to other services [36].

Real-time logging and anomaly queues allowed tax officers to prioritize intervention within seconds of detection, affirming

the system's applicability in time-sensitive revenue protection operations [37].

Overall, the stress tests validated the system's robustness under live data conditions, making it suitable for deployment in dynamic tax environments with fluctuating transaction volumes and complex reporting timelines [38].

7.3 Known Limitations and Future Areas for Improvement

Despite its strong performance, the proposed AI-driven tax surveillance system faces several limitations that must be addressed for broader adoption and sustained efficacy [39]. One primary concern is the model's dependency on high-quality labeled data. In many developing countries, historical fraud case data is sparse, unstructured, or inconsistently labeled, limiting the ability to train supervised models effectively [40].

Additionally, while the current system performs well on structured transactional data, it remains less effective in detecting fraud embedded in semi-structured or unstructured sources such as scanned invoices, emails, or legal contracts. Integrating natural language processing (NLP) and document classification tools would expand the system's coverage into these domains [41].

The model's interpretability also poses challenges. While SHAP values provide some level of explanation for decisions, more intuitive, user-friendly interfaces are needed to help non-technical tax officers understand why certain taxpayers were flagged. This is crucial for ensuring procedural fairness and defending decisions during appeals or legal proceedings [42].

Ethical concerns surrounding algorithmic bias must also be addressed. If models are trained on biased or unbalanced datasets, they may unfairly target certain demographic or economic groups. Implementing fairness-aware algorithms and continuous bias audits is essential for equitable enforcement [43].

Finally, the governance of adaptive learning poses regulatory challenges. As models evolve through retraining, it becomes harder to maintain consistent auditability and compliance documentation. A version-controlled model registry and regulatory sandbox environment are recommended for oversight [44].

Future improvements include expanding the system's multilingual capabilities, enhancing integrations with biometric tax ID systems, and developing federated learning techniques for secure model training across jurisdictions [45].

By acknowledging and addressing these limitations, the system can evolve into a global standard for proactive, transparent, and data-driven tax administration in the digital age [46].

8. IMPLEMENTATION STRATEGY AND POLICY FRAMEWORK

8.1 Phased Deployment Across Tax Jurisdictions

Deploying a cloud-based AI surveillance system for tax fraud detection requires a phased implementation approach, tailored to the administrative, technological, and regulatory capacities of each tax jurisdiction. This ensures that rollout is efficient, minimally disruptive, and adaptable to context-specific needs [30].

Phase one involves pilot testing in a limited, high-risk sector—such as VAT fraud in the e-commerce industry or transfer pricing violations in corporate tax. The system is deployed in parallel with existing manual audit workflows, allowing comparative validation and stakeholder acclimatization [31]. Pilot regions are selected based on available digital infrastructure, data maturity, and prior leakage patterns.

Phase two focuses on scaling horizontally across sectors, such as integrating customs, excise, and property tax data. This phase includes deploying secure APIs for cross-agency data exchange and integrating feedback loops from enforcement actions back into the model training environment [32]. Cloud resources are scaled as needed using infrastructure-as-code templates to ensure uniform deployment.

Phase three is jurisdictional expansion. The system is rolled out nationally or across federal states, with customization for regional tax regulations, languages, and user interfaces. At this stage, AI risk engines are re-tuned to accommodate regional transaction norms, business behaviors, and localized fraud typologies [33].

Final deployment phase includes cross-border coordination. Jurisdictions align on shared tax standards and data governance protocols to support multinational tax enforcement and joint audits. Interoperability with OECD's Common Reporting Standard (CRS) and EU's DAC7 reporting framework enables cross-jurisdictional visibility of income, assets, and digital sales [34].

Each phase includes continuous monitoring, stakeholder feedback collection, and iterative system updates. This approach mitigates risks associated with "big bang" implementations and allows tax authorities to build public confidence in the technology while gradually modernizing their compliance infrastructure [35].

8.2 Training, Interagency Collaboration, and Stakeholder Buy-In

The successful adoption of an AI-powered tax surveillance system hinges not only on its technical deployment but also on the training and alignment of stakeholders. Governments must invest in building internal capacity and fostering interagency collaboration to fully realize the system's benefits [36].

Training programs should be developed for tax officers, auditors, legal experts, and IT personnel. These programs must cover core topics such as AI literacy, interpreting machine learning outputs, ethical use of predictive analytics, and interacting with model dashboards. Tiered training—beginner to advanced—ensures that personnel at all levels are equipped to leverage the system appropriately [37].

Beyond technical skills, it is essential to cultivate critical understanding of the system's limitations. Tax officers must learn to contextualize alerts, exercise human judgment, and communicate findings transparently during audits or legal proceedings. Scenario-based simulations using synthetic data can reinforce practical application [38].

Interagency collaboration is equally important. Revenue authorities, customs departments, financial intelligence units, and ministries of justice must coordinate workflows, establish data-sharing agreements, and agree on escalation protocols for high-risk cases. Shared case management platforms and common data dictionaries can reduce duplication and foster consistency in investigations [39].

Securing **stakeholder buy-in**—especially from policymakers, unions, and the public—requires transparent communication about the system's goals, benefits, and guardrails. Publishing periodic performance reports, success stories, and oversight committee reviews can enhance legitimacy [40]. Involving external experts and civil society in governance forums adds an additional layer of accountability and public trust.

Moreover, tax administrations should collaborate with international organizations such as the World Bank, IMF, and OECD to align with global best practices and access technical assistance. These alliances reinforce credibility and facilitate knowledge transfer across jurisdictions [41].

8.3 Legal, Ethical, and Data Governance Considerations

Introducing AI into tax administration requires careful navigation of legal frameworks, ethical principles, and data governance standards to ensure the system operates within public interest boundaries [42].

From a legal perspective, AI deployment must align with data protection laws and taxpayer rights legislation. This includes compliance with local data privacy acts, cross-border data transfer regulations, and international human rights conventions. Agencies must secure explicit legal mandates to process sensitive financial data through AI algorithms and store it in cloud environments [43].

Ethically, tax surveillance must avoid algorithmic discrimination, overreach, and opaque decision-making. To address this, models should be subjected to fairness audits, and their outputs must be explainable to both auditors and affected taxpayers. Embedding transparency features such as traceable audit trails, dispute resolution workflows, and bias-mitigation protocols is vital [44].

Data governance underpins the system's legitimacy. Clear policies must define data ownership, access rights, retention periods, and classification levels. Agencies should implement data stewardship roles to oversee compliance and ensure proper custodianship across departments [45]. Public sector use of AI also requires accountability mechanisms—such as ethics boards, parliamentary review, and whistleblower protections—to prevent misuse or abuse of predictive enforcement powers.

Cloud contracts with third-party vendors must include sovereignty clauses ensuring that data remains within agreed national jurisdictions. Encryption, anonymization, and synthetic data testing further protect citizen privacy while supporting model development.

Ultimately, embedding ethical AI principles and robust data governance into system design ensures that technological innovation serves public revenue goals without compromising democratic norms or citizen rights [46].

9. FUTURE OUTLOOK AND SCALABILITY ACROSS BORDERS

9.1 Harmonizing AI Tax Systems Across Nations

In an increasingly interconnected global economy, harmonizing AI-driven tax systems across nations is essential to combat transnational tax evasion and enhance collective revenue resilience [35]. While many governments have developed AI surveillance tools independently, the absence of shared standards, communication protocols, and regulatory alignment limits the effectiveness of these efforts, particularly in cases involving multinational corporations and digital platform economies [36].

To promote interoperability, countries must align their data schemas, fraud typologies, and enforcement procedures. Adoption of common formats such as the OECD's Standard Audit File for Tax (SAF-T) and ISO 20022 for financial messaging can facilitate seamless exchange of tax intelligence [37]. Furthermore, developing AI interoperability frameworks that allow national systems to interact securely—without exposing sensitive data—will be crucial for collective enforcement action.

Harmonization also requires policy convergence. Countries need to adopt compatible AI governance strategies, including standards for algorithm transparency, model auditability, and bias mitigation. Regional coalitions like the African Tax Administration Forum (ATAF) or EU's Fiscalis program can play a leading role by offering shared infrastructures and training platforms [38].

To avoid duplicative effort and foster cost-sharing, nations may pool resources for developing open-source fraud detection models, supported by multilateral institutions such as the IMF and World Bank. This collaboration not only enhances the technical robustness of tax AI systems but also

ensures smaller economies are not left behind in digital enforcement evolution [39].

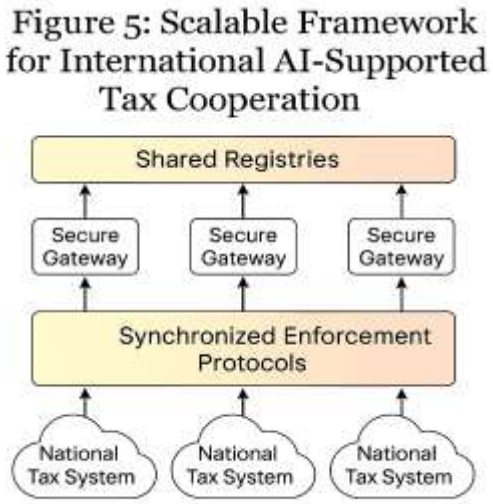

Figure 5: Scalable Framework for International AI-Supported Tax Cooperation illustrates how national systems can federate through secure gateways, shared registries, and synchronized enforcement protocols.

9.2 AI Ethics, Cross-Border Data Exchange, and Global Tax Justice

While AI technologies hold transformative potential for tax enforcement, their cross-border application raises significant concerns regarding ethics, sovereignty, and fairness. To support a just global tax system, international cooperation on AI ethics and data governance must be prioritized alongside technological standardization [40].

Cross-border data exchange, essential for tracking income and assets hidden across jurisdictions, is constrained by inconsistencies in privacy legislation, lack of mutual legal frameworks, and geopolitical tensions. Agreements like the OECD's Common Reporting Standard (CRS) have laid the foundation for information sharing, but their integration with AI surveillance tools remains minimal and fragmented [41]. A unified approach is needed to ensure AI systems can securely consume and act upon foreign-sourced data without violating privacy rights or national laws.

To facilitate ethical AI deployment, global bodies such as the UN, OECD, and G20 should establish a Tax AI Governance Charter outlining principles of transparency, non-discrimination, auditability, and proportionality. This charter could serve as a benchmark for AI system design and operational safeguards in revenue administrations [42].

Data minimization principles must also guide cross-border analytics. AI models should only process information directly relevant to tax risk assessment, and all cross-jurisdictional data access must be logged and monitored. Further, differential privacy techniques, federated learning, and secure multiparty computation (SMPC) can enable collaborative analytics without revealing raw taxpayer data [43].

Beyond technical safeguards, ethical AI in tax must advance global tax justice. The system should prioritize detection of corporate tax abuse and illicit flows rather than overly targeting individuals from lower-income brackets. Ensuring that enforcement is balanced, inclusive, and protective of human rights is central to avoiding reputational damage and fostering international trust [44].

Achieving this balance requires multi-stakeholder oversight, including civil society, digital rights groups, and independent regulators. International forums must be empowered to audit AI system performance, review algorithmic impacts, and mediate disputes arising from automated enforcement decisions [45].

Ultimately, the responsible global deployment of AI in taxation hinges not only on interoperability and performance, but also on ethical alignment, legal safeguards, and a shared vision of equitable taxation in the digital age [46].

10. CONCLUSION AND RECOMMENDATIONS

10.1 Summary of Contributions, Practical Takeaways, and the Road Ahead

This paper presents a comprehensive framework for leveraging cloud-based artificial intelligence (AI) to enhance the predictive detection of tax-related financial crimes. In an era where digitalization has introduced both unprecedented opportunities and risks for tax administrations, the integration of intelligent technologies into compliance systems marks a pivotal evolution in public revenue management.

The primary contribution of this work lies in the design and evaluation of a modular AI-driven tax surveillance architecture that combines real-time anomaly detection, machine learning pipelines, and secure data infrastructure. The model facilitates seamless data ingestion from multiple sources, including e-filing systems, transactional logs, and third-party databases, while preserving privacy and ensuring regulatory compliance through encryption and role-based access controls.

Key insights emerge from the case simulation of VAT and corporate tax evasion, where the AI system demonstrated significant advantages over traditional audit methods. The model exhibited faster detection times, higher fraud classification accuracy, and improved resource allocation by automating risk scoring and audit prioritization. These improvements translate directly into measurable gains in revenue recovery, institutional efficiency, and taxpayer compliance.

In practical terms, this study outlines a clear roadmap for phased deployment of AI systems across tax jurisdictions. By starting with pilot programs in high-risk sectors, governments

can validate performance before expanding horizontally across departments and vertically across regional and national levels. Training tax officers, fostering interagency collaboration, and engaging stakeholders early are essential steps for successful adoption.

Furthermore, the importance of ethical governance and data stewardship is underscored throughout the study. As tax authorities adopt algorithmic decision-making, they must ensure transparency, auditability, and fairness. Establishing internal oversight bodies, adhering to global AI ethics guidelines, and applying data minimization principles will help mitigate the risks of misuse and bias. This approach not only safeguards civil liberties but also builds public trust—an indispensable asset in tax administration.

From a policy perspective, the paper calls for increased international cooperation to harmonize AI-driven tax enforcement systems. By adopting common technical standards and interoperability frameworks, countries can coordinate responses to transnational tax evasion and reduce compliance gaps. Shared infrastructure and open-source model development offer viable pathways for resource-constrained jurisdictions to participate in global tax intelligence networks.

Looking ahead, the road to widespread implementation involves continuous model refinement, legal adaptation, and infrastructure scaling. Advancements in federated learning, privacy-preserving computation, and explainable AI will further enhance the capabilities and acceptability of these systems. Investment in human capital—through training and inclusive digital transformation—will be equally important to ensure sustainable impact.

In conclusion, the integration of cloud-based AI into tax systems represents a transformative step toward smarter, fairer, and more responsive public finance. By embracing this technology with caution and accountability, governments can modernize their enforcement strategies, close the tax gap, and reinforce fiscal sovereignty in an increasingly complex economic landscape.

11. REFERENCE

1. Bezditnyi V. Use of Artificial Intelligence for Tax Planning Optimization and Regulatory Compliance. Research Corridor Journal of Engineering Science. 2024 Jan 10;1(1):103-42.
2. ESCAP U. The digitalization of tax administrations in Asia and the Pacific: a manual for practitioners.
3. Wolf S. Does aid improve public service delivery?. Review of world economics. 2007 Dec;143:650-72.
4. Adeniji EH. Leveraging enterprise analytics to align risk mitigation, health IT deployment, and continuous clinical process improvement. *International Journal of Science and Research Archive*. 2023;10(2):1314–1329. doi: https://doi.org/10.30574/ijsra.2023.10.2.1003.
5. Reinikka R, Svensson J. Survey techniques to measure and explain corruption. World Bank Publications; 2003.
6. Noah GU. Interdisciplinary strategies for integrating oral health in national immune and inflammatory disease control programs. *Int J Comput Appl Technol Res*. 2022;11(12):483-498. doi:10.7753/IJCATR1112.1016.
7. Wong C. Rebuilding government for the 21st century: can China incrementally reform the public sector?. The China Quarterly. 2009 Dec;200:929-52.
8. Azfar O, Kahkonen S, Lanyi A, Meagher P, Rutherford D. Decentralization, governance and public services: The impact of institutional arrangements. InDevolution and development 2018 Jan 18 (pp. 45-88). Routledge.
9. Chukwunweike Joseph, Salaudeen Habeeb Dolapo. Advanced Computational Methods for Optimizing Mechanical Systems in Modern Engineering Management Practices. *International Journal of Research Publication and Reviews*. 2025 Mar;6(3):8533-8548. Available from: https://ijrpr.com/uploads/V6ISSUE3/IJRPR40901.pdf
10. Bachner G, Bednar-Friedl B. The effects of climate change impacts on public budgets and implications of fiscal counterbalancing instruments. Environmental Modeling & Assessment. 2019 Apr 1;24:121-42.
11. Haque NU, Sahay R. Do government wage cuts close budget deficits? Costs of corruption. Staff Papers. 1996 Dec 1;43(4):754-78.
12. Reinikka R, Svensson J. Explaining leakage of public funds. WIDER Discussion Paper; 2001.
13. Oates WE. "Automatic" Increases in Tax Revenues—The Effect on the Size of the Public Budget. InFinancing the new federalism 2015 Sep 25 (pp. 139-160). Routledge.
14. Downes TA, Figlio DN. Do tax and expenditure limits provide a free lunch? Evidence on the link between limits and public sector service quality. National Tax Journal. 1999 Mar 1;52(1):113-28.
15. Joyce PG, Mullins DR. The changing fiscal structure of the state and local public sector: The impact of tax and expenditure limitations. Public administration review. 1991 May 1:240-53.
16. Emi-Johnson Oluwabukola, Fasanya Oluwafunmibi, Adeniyi Ayodele. Predictive crop protection using machine learning: A scalable framework for U.S. Agriculture. Int J Sci Res Arch. 2024;15(01):670-688. Available from: https://doi.org/10.30574/ijsra.2024.12.2.1536
17. Ahmed V, Nazir A, Gregory D, Faraz Z, Ace T. Issue paper-social enterprise development in Pakistan: the way forward.
18. Wood T, Basto-Fernandes V, Boiten E, Yevseyeva I. Systematic Literature Review: Anti-Phishing Defences and Their Application to Before-the-click Phishing Email Detection. arXiv preprint arXiv:2204.13054. 2022 Apr 27.
19. Laidlaw E. Privacy and cybersecurity in digital trade: The challenge of cross border data flows. Available at SSRN 3790936. 2021 Feb 22.

20. Emi-Johnson Oluwabukola, Nkrumah Kwame, Folasole Adetayo, Amusa Tope Kolade. Optimizing machine learning for imbalanced classification: Applications in U.S. healthcare, finance, and security. Int J Eng Technol Res Manag. 2023 Nov;7(11):89. Available from: https://doi.org/10.5281/zenodo.15188490
21. Kurt AC. Auditor Expertise in Government Contracting. Available at SSRN 4199751. 2022 Aug 25.
22. Oladipupo AO. A smarter path to growth: why SMEs need FP&A and M&A strategies to compete in a global economy. *Int J Comput Appl Technol Res*. 2022;11(10):1–12. doi:10.7753/IJCATR1110.1001.
23. Risse R, Lang M. Tax Law and Digitization.
24. Chukwunweike J, Lawal OA, Arogundade JB, Alade B. Navigating ethical challenges of explainable AI in autonomous systems. *International Journal of Science and Research Archive*. 2024;13(1):1807–19. doi:10.30574/ijsra.2024.13.1.1872. Available from: https://doi.org/10.30574/ijsra.2024.13.1.1872.
25. Tian GY, McFarland T, Guo S. Automated decision making and deportation: legal concerns and regulation. Griffith Law Review. 2025 Mar 18:1-28.
26. Olayinka OH. Big data integration and real-time analytics for enhancing operational efficiency and market responsiveness. *Int J Sci Res Arch*. 2021;4(1):280–96. Available from: https://doi.org/10.30574/ijsra.2021.4.1.0179
27. Kababiito Lillian. Harnessing Artificial Intelligence for Real-Time Compliance in the U.S. Oil & Gas Sector: Enhancing Tax Accuracy, Curbing Evasion, and Unlocking Revenue Growth through Intelligent Automation. *International Journal of Computer Applications Technology and Research*. 2025;14(05):55–70. doi:10.7753/IJCATR1405.1006.
28. Devi S. Business Documentation: A Technical Communication Skill. Chyren Publication; 2025 Apr 2.
29. Olayinka OH. Ethical implications and governance of AI models in business analytics and data science applications. *International Journal of Engineering Technology Research & Management*. 2022 Nov;6(11). doi: https://doi.org/10.5281/zenodo.15095979.
30. Owens J, Risse R, editors. Tax Law and Digitalization: The New Frontier for Government and Business: Principles, Use Cases and Outlook. Kluwer Law International BV; 2021 Sep 15.
31. Lukianykhina O, Suprunenko S, Slavkova A, Skorba O, Zavrazhnyi K. Digital Innovations in the Fiscal Policy of Ukraine: Promoting Sustainable Economic Development. International Journal of Economics and Financial Issues. 2024 Jul 3;14(4):77-86.
32. Gandhi H, Tandon K, Gite S, Pradhan B, Alamri A. Navigating the complexity of money laundering: anti–money laundering advancements with AI/ML insights. International Journal on Smart Sensing and Intelligent Systems. 2024(1).
33. Rahayu SK, Kusdianto A. Challenges of digital tax administration transformation in Indonesia. InBusiness and Management Annual Volume 2023 2023 Jun 19. IntechOpen.
34. Devi S. Corporate Taxation. Chyren Publication; 2025 Apr 2.
35. Song J. AI of Public Communication for the Vulnerable: A Focus on Voice-Based Chatbots and Policy Suggestions. InAdvanced Virtual Assistants-A Window to the Virtual Future 2023 Sep 14. IntechOpen.
36. Song J. The implications of providing voice-based chatbots in public service for digital inclusion and public communication. OMNES: The Journal of Multicultural Society. 2022 Jul;12(2):26-68.
37. Omopariola BJ, Aboaba V. Advancing financial stability: The role of AI-driven risk assessments in mitigating market uncertainty. *International Journal of Scientific Research and Advances*. 2021 Sep;3(2). doi: 10.30574/ijsra.2021.3.2.0106
38. Olayinka OH. Data driven customer segmentation and personalization strategies in modern business intelligence frameworks. *World Journal of Advanced Research and Reviews*. 2021;12(3):711–726. doi: https://doi.org/10.30574/wjarr.2021.12.3.0658.
39. Yayman D. Taxation in Virtual Worlds: Analysis Under United States of America and Turkish Tax Regulations. Sosyoekonomi. 2023 Jan 1;31(55):211-31.
40. Brondolo J, Brondolo MJ, Chooi A, Schloss T, Siouclis A. Compliance risk management: developing compliance improvement plans. International Monetary Fund; 2022 Mar 18.
41. Song J. The Introducing voice-based public services for strengthening the accessibility of the social vulnerables and open public communication. Journal of Intelligence and Information Systems. 2022;28(2):279-306.
42. Mints A, Sidelov P. Digital payment card fraud: new vectors and detection. Digital Technologies in the Contemporary Economy: Collective Monograph/editor Žaneta Simanavičienė. ISBN 9786094880506. 2022.
43. Pender K, Cherkasova S, Yamaoka-Enkerlin A. Compliance and whistleblowing: How technology will replace, empower and change whistleblowers. InFinTech 2024 May 21 (pp. 485-522). Edward Elgar Publishing.
44. Omopariola BJ. Decentralized energy investment: Leveraging public-private partnerships and digital financial instruments to overcome grid instability in the U.S. *World Journal of Advanced Research and Reviews*. 2023 Dec;20(03):2178–2196. doi: 10.30574/wjarr.2023.20.3.2518
45. Busari TA, Dada SO, Ajala MO. TAXES FROM INCOME, PROFIT AND CAPITAL GAINS AMONG SUB-SAHARA AFRICAN COUNTRIES: IMPACT OF INDUSTRIALIZATION. Fiscal.
46. Jones N. The AZ of Payments: A Modern and Practical Glossary for the Curious and the Forgetful. Taylor & Francis; 2025 Mar 14.