



Machine learning-based risk analytics in combating the shadow economy: public indicators on Kaşif and the debate on procedural safeguards

Salahattin Altundağ

Assistant Professor

Dicle University

Türkiye

E-mail: salahattinaltundag@gmail.com; ORCID: 0000-0002-6198-7959

Keywords

Algorithmic Governance, KAŞIF, KURGAN, Tax Auditing, Shadow Economy, Explainable Artificial Intelligence (XAI), Fraudulent Documentation, Document Analysis, Human-in-the-Loop

Abstract

In efforts to combat the shadow economy and the use of fraudulent documentation, traditional auditing systems remain inadequate in the face of massive data volumes. This study aims to evaluate the implications of the machine learning-based KAŞIF and KURGAN systems, implemented by the Ministry of Treasury and Finance, for tax auditing. The study employs descriptive document analysis based on official statistics and bulletins. The compiled public indicators suggest that algorithmic governance narrows the time asymmetry between the commission and detection of the offence. The documents report that the system analyzes 4 million taxpayers per day; that, in illustrative cases, a fraudulent invoice network worth ₺246 million was detected; and that a transaction volume of ₺276 billion was reported as “risky” over the last three months. However, turning opaque autonomous risk scoring into administrative sanctions entails procedural risks in relation to the presumption of innocence, proportionality, and legal certainty. Consequently, for the sustainability of algorithmic auditing, the study recommends the adoption of Explainable Artificial Intelligence (XAI) standards and the integration of a human-in-the-loop model into administrative procedures, where final decisions are left to expert auditors.

Citation

Altundağ, S. (2026). Machine learning-based risk analytics in combating the shadow economy: public indicators on Kaşif and the debate on procedural safeguards. *Science, Education and Innovations in the Context of Modern Problems*, 9(6), 1–10. <https://doi.org/10.56334/sei/9.6.21>

Licensed

© 2026 The Author(s). Published by *Science, Education and Innovations in the Context of Modern Problems (SEI)*, under the auspices of IMCRA – International Meetings and Conferences Research Association (Azerbaijan).

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

<http://creativecommons.org/licenses/by/4.0/>

Received: January 12, 2025

Accepted: March 24, 2026

Published Online : May 29, 2026

1. INTRODUCTION

The shadow economy, one of the leading structural problems of public finance, and the use of “fraudulent invoices”—commonly referred to in Turkish as *naylon fatura* and defined in the literature as false or misleading documents (SMİYB—*Sabte veya Muhteviyatı İtibarıyla Yanıltıcı Belge*)—together constitute a key factor that erodes the tax base and threatens macroeconomic stability. Informal economic activities not only undermine the state’s tax revenues but also create an asymmetric environment of unfair competition to the detriment of compliant taxpayers in the market (Çember, 2026). Although tax administrations have historically relied on audit mechanisms to address this macroeconomic problem, traditional audit methods have increasingly become inadequate in the face of a digitized and progressively more complex commercial ecosystem.

Traditional tax auditing practices are often based on reactive examinations and audits conducted after tax returns have been filed. Moreover, since comprehensive audits may cover multiple taxes, issues, and tax years, detection and administrative action may take place considerably later than the period in which the taxable event occurred (OECD, 2017, pp. 19, 58). This retrospective structure indicates that administrations may have to address even “old” risks and fraudulent practices that occurred years earlier. It may therefore

create a time-lag channel that enables short-lived or shell entities to magnify public losses and exit the system (OECD, 2017, p. 52). Earlier systems such as the Fraudulent Document Risk Analysis Program (SARP) and the Risky Refund Monitoring and Analysis Program (RİTAP) could experience flexibility problems in responding to rapidly changing fraud typologies because they operated on the basis of predefined rules (rule-based systems) (Anadolu Ajansı, 2026; Çember, 2026). On the other hand, particularly in high-volume declaration and refund processes such as VAT, the frequent infeasibility of direct expert examination has led administrations to turn to new technologies and advanced analytics for risk detection and prioritization. Accordingly, approaches such as machine learning and anomaly detection have emerged as both an academic and practical necessity (OECD, 2016, p. 21; OECD, 2017, p. 33; Yıldız & Köşer, 2025, p. 9).

The main objective of this study is to systematically compile, through document analysis, the publicly disclosed detection indicators concerning the KAŞİF and KURGAN systems, which were implemented in cooperation between the Revenue Administration of the Ministry of Treasury and Finance (GİB) and the TÜBİTAK Artificial Intelligence Institute. The study also discusses what procedural safeguard needs these indicators point to in terms of evaluability and accountability (Uyumsoft, 2026). By focusing on data shared by public authorities regarding the daily screening of approximately 4 million taxpayers through terabyte-scale big data platforms, the study aims to present the implications of algorithmic governance for public finance within a descriptive framework (Ekonomi Gazetesi, 2026).

However, the scope of this research is not strictly limited to an analysis of technological efficiency. One of its central aims is also to discuss the socio-legal risks that may arise when algorithmic risk scores assigned by the system are transformed into administrative sanctions directly affecting the freedom to conduct business, such as the ex officio termination of taxpayer registration.

The subsequent sections of this study are structured as follows. First, the conceptual foundation of algorithmic auditing is presented, and KAŞİF's machine learning-based operation is described to the extent permitted by publicly available statements. Second, the research design and document analysis approach are explained. Public indicators systematically extracted from the documents—such as figures reported as risky transaction volumes and illustrative cases concerning relationship networks and network structures—are then presented within a descriptive framework. Finally, the discussion section evaluates the procedural tensions that may arise from the use of algorithmic risk classifications in administrative processes in relation to the principles of the presumption of innocence, proportionality, and legal certainty. It then develops policy recommendations along the axes of explainability (XAI) and the human-in-the-loop model.

2. LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

This section offers a conceptual analysis of algorithmic governance practices in tax auditing with reference to the OECD's recent vision documents and the main behavioral models in the literature on tax psychology.

2.1. Algorithmic Governance in Public Finance and Tax Administration 3.0

Digitalization in public finance has evolved beyond a mere instrument of automation into autonomous decision-support mechanisms based on machine learning and deep learning. The "Tax Administration 3.0" vision conceptualized by the OECD represents a new paradigm in which tax administrations minimize human intervention and incorporate digital identity and advanced analytics systems into tax governance (Yıldız & Köşer, 2025, p. 14). This paradigm envisages the collection of taxpayer data in big data pools and its real-time analysis through complex algorithms. Indeed, the OECD's Tax Administration 3.0 vision presents the autonomization of processes as a global standard, whereby tax administrations operate embedded within "natural systems" through digital identity and advanced analytics (OECD, 2020, p. 11).

An examination of the digital maturity levels of tax authorities worldwide reveals that multilayer perceptron artificial neural networks are reported to be actively used for the detection of tax evasion. For example, the literature reports that the artificial neural network models implemented by the Spanish Tax Administration achieved a success rate of 84.3% (Yıldız & Köşer, 2025, p. 16). Türkiye's algorithmic auditing maturity, aligned with the Tax Administration 3.0 vision, has reached a globally competitive level by leveraging massive data obtained from e-Invoice, e-Ledger, and digital tax return systems and incorporating them into rule-independent, predictive artificial intelligence architectures that go beyond simple electronic screening.

2.2. The Slippery Slope Framework in Taxation

The structural effects of algorithmic governance and autonomous auditing systems on taxpayer behavior can be explained within the framework of the Slippery Slope Framework developed by Erich Kirchler, Erik Hoelzl, and Ingrid Wahl (2008). The theory identifies perceived audit power and trust in the administration as the two main axes determining tax compliance. As audit power increases, compliance is supported more strongly through the channel of enforced compliance; as the level of trust rises, the voluntary compliance component becomes stronger. The highest and most sustainable level of compliance is expected when both elements are high (Kirchler et al., 2008; Özmumcu, 2026). This framework enables a simultaneous evaluation of the risk that artificial intelligence-

supported audit architectures may increase “audit power” while eroding the “trust” channel if explainability and procedural safeguards remain weak.

According to the literature, scenarios in which high trust and strong audit power coexist generate an ideal synergy between voluntary and enforced tax compliance. The artificial intelligence-supported audit mechanisms implemented in Türkiye fundamentally transform and maximize the administration’s “audit power.” However, the most critical risk identified by the theory is that artificial intelligence algorithms may operate as closed “black boxes” and produce potential false positive results, thereby impairing procedural justice and eroding taxpayers’ trust (Özmumcu, 2026). This risk of trust loss, triggered by a “perception of excessive power,” makes the transition to Explainable Artificial Intelligence (XAI) models a theoretical necessity, as such models would help ensure the transparency and accountability of risk scoring systems (Özmumcu, 2026). The “power–trust” asymmetry articulated by the Slippery Slope Framework provides the technical justification for the transition in tax auditing from static and reactive systems to more intelligent, flexible, and accurate autonomous systems.

3. KAŞİF’S TECHNICAL ARCHITECTURE AND RISK-SCORING ALGORITHMS

To mitigate the structural risks identified in the literature on tax psychology and to place the administration’s audit power within an efficient algorithmic framework, Türkiye has made a radical technological leap in its tax auditing architecture.

3.1. Transition from Rule-Based Systems to Autonomous Machine Learning

Historically, systems used in Türkiye to combat the shadow economy and fraudulent documentation—such as the Fraudulent Document Risk Analysis Program (SARP) and the Risky Refund Monitoring and Analysis Program (RİTAP)—were structurally based on static, rule-based mechanisms (Anadolu Ajansı, 2026; Ekonomi Gazetesi, 2026). The most fundamental limitation of these older-generation systems was their dependence on predefined scenarios and rigid threshold values, which prevented them from responding immediately to new and complex fraud typologies. The ability of malicious actors to learn these static rules and circumvent parameter boundaries limited the proactive audit capacity of such systems.

To overcome these limitations, the Establishment-Supervised Analysis system, known as KURGAN, was first implemented within the Tax Inspection Board (VDK). Operating on the logic of big data processing, KURGAN performs relationship-network analysis based on 13 different objective criteria and represents one of the administration’s first major steps toward proactive auditing (Softa, 2025, pp. 11–12). The culmination of this digital transformation, however, is KAŞİF, developed by the Revenue Administration in cooperation with the TÜBİTAK Artificial Intelligence Institute. According to public disclosures, KAŞİF is presented as a machine learning-based risk analytics application that does not operate on a rule-based logic; instead, it analyzes very large datasets daily without relying on prior definitions or rule sets and flags risky situations. However, details concerning the model’s learning type—such as whether it is supervised or unsupervised—and its validation protocols are not publicly available. For this reason, the technical classification in this study is constructed cautiously, only to the extent permitted by public statements (Palabıyık & Uylaş, 2026; Anadolu Ajansı, 2026). The system performs automatic risk scoring by learning unusual behavioral deviations in the data itself, thereby eliminating the cumbersome structure of rule-based systems (Ekonomi Gazetesi, 2026).

3.2. Variables Analyzed and Data Processing Capacity

The effectiveness of KAŞİF’s risk-scoring algorithms is based on the scale of its data mining capacity. Rather than relying on one-dimensional audits, the machine learning model conducts a multivariate econometric analysis that captures inconsistencies among the financial, physical, and legal realities of businesses.

In terms of its technical architecture, KAŞİF has the capacity to screen approximately 4 million taxpayers per day in real time through massive big data platforms comprising tens of terabytes (Anadolu Ajansı, 2026). The extensive dataset forming the input matrix of the system’s algorithms includes taxpayer registration and registry information, e-invoice and document traffic, tax accrual and collection ratios, import and export data, submitted tax returns and notifications, employee counts based on Social Security Institution (SGK) records, business and production capacity data, and inventories of tangible fixed assets recorded among assets (Ekonomi Gazetesi, 2026). KAŞİF models a company’s declared massive transaction volumes as output and its physical and employment-based production capacity as input within a simultaneous cross-checking function. It thereby classifies deviations from “business as usual”—that is, from the natural flow of commercial life—as structural anomalies and calculates the algorithmic risk score within seconds.

4. RESEARCH METHODOLOGY

This section details the research design, data collection techniques, and analytical limitations developed to examine the implications of the KAŞİF application, implemented by the Ministry of Treasury and Finance, for tax auditing.

4.1. Research Model and Data Collection Process

This research is designed as a document analysis study based on a descriptive approach. The main dataset of the study consists of official statistics, institutional reports, and press agency bulletins made publicly available by the Ministry of Treasury and Finance and the Revenue Administration (GİB) (Ekonomi Gazetesi, 2026). The analysis of these relevant documents relies on the systematic extraction and compilation of publicly available performance indicators concerning variables generated through KAŞİF's data mining processes, such as accrual, collection, imports, exports, business capacity, personnel employment, and e-document traffic (Palabiyik & Uylaş, 2026).

4.1.1. Document Corpus and Selection Criteria

The document corpus examined within the scope of this descriptive document analysis consists of official reports published by the Ministry of Treasury and Finance, institutional bulletins, reports containing statistical data issued by national news agencies such as Anadolu Agency (AA), and expert assessments related to the subject.

The following criteria were applied in the selection process:

- **Time Period:** The period from the last quarter of 2025 to the first quarter of 2026, during which the operational results of the KAŞİF system began to be reflected in the public domain.
- **Inclusion Criteria:** Texts containing direct “numerical indicators or concrete illustrative cases” regarding the system’s technical capacity or audit outcomes were included in the analysis.
- **Exclusion Criteria:** Speculative opinion columns not based on data and unofficial blog posts expressing only personal views were excluded from the main analytical universe in accordance with the evidence hierarchy. Such secondary sources, for example Aslanpınar (2026), were included in the discussion only in the context of “field perspective and professional commentary.”

4.1.2. Coding Scheme and Data Extraction

In order to subject the documents to systematic analysis and to move beyond the limits of a mere “news compilation,” the data were classified through a specific coding scheme, or data extraction form:

- **Type of Indicator:** Detected transaction volume, identified network or case, speed/process optimization, and accuracy claim.
- **Source Class:** Primary/official statements with high evidentiary value; agency bulletins with moderate evidentiary value; media/blog/expert commentaries with low evidentiary value or field-perspective relevance.
- **Mode of Expression:** Officially “reported,” “claimed” by experts, or presented to the public as an “illustrative case.”
- **Interpretive Boundary (Analytical Caveat):** The concept of “detection” was not treated as absolutely equivalent to “prevention,” and the caveat that publicly disclosed examples do not necessarily indicate the system’s “general or independently verified performance” was maintained consistently across all findings. This interpretive boundary was applied in line with the logic of the READ (Ready–Extract–Analyse–Distil) approach, which separates data from inference and interprets findings cautiously according to evidentiary strength (Dalglish, Khalid, & McMahon, 2020, p. 1424).

4.2. Indicators and Descriptive Evaluation Framework

This study does not constitute a causality test and does not aim to provide experimental or quasi-experimental proof based on dependent and independent variables. The cautious approach adopted here is a deliberate methodological limitation designed not merely to address administrative data constraints, but also to manage the difficulty of reliably measuring performance and impact without an independent control group, because audit data are endogenously selected, a problem frequently discussed in the literature as selective labels (Battaglini et al., 2022, p. 2). Accordingly, the “results” of the transition from traditional rule-based audit mechanisms to algorithmic governance and risk scoring are not treated here as independently verified performance metrics. Rather, they are examined within a descriptive evaluation framework in which indicators shared by public authorities are systematically compiled.

Within this framework, “public indicators” are defined not as a single-layered performance measure, but as a set of analytical signals stratified across four levels:

- I. **Capacity/input indicators:** Statements describing the system’s screening capacity, such as the number of taxpayers analyzed daily.
- II. **Algorithmic prioritization/flagging indicators:** Classifications such as those indicating that a taxpayer is considered to be at “high risk.”

III. Administrative process triggers: Workflow signals that initiate verification gateways, such as referring risky findings to audit/compliance processes or conducting on-site inspections.

IV. Procedural/administrative outcomes: Administrative decisions established in light of concrete evidence.

Among these layers, the category of “flagging/risk” at level II and the category of “legal outcome” at level IV are not equivalent. Expressions such as “detected” or “reported” refer to the level of finding disclosure announced by the public authority and do not, by themselves, imply judicial finality or a definitive legal characterization. Therefore, in this study, references to “detection” are treated not as verified violations or final sanctions, but as publicly disclosed findings.

In accordance with this methodological framework, the main indicators analyzed are descriptively classified under the following headings: (a) “the number of risky or shell taxpayers reported as detected or as having a high level of risk,” (b) “the volume of fraudulent document traffic reported as detected within a time series,” and (c) “unlawful VAT refund attempts reported as a result of algorithmic analysis” (Uyumsoft, 2026; Palabiyık & Uylaş, 2026).

4.3. Limitations of the Dataset

The most fundamental factor defining the methodological boundaries of this study is that the technological infrastructure under analysis is strictly protected within the framework of the state’s fiscal security and the principle of “tax confidentiality.” In this study, direct access could not be obtained to the specific decision trees, validation protocols, or raw time-series data of the algorithmic architecture operated by GİB. Instead, inferences were made only on the basis of data consolidated by the administration and shared with the public. Moreover, detailed statistics on false positive rates, which are frequently discussed in the literature on artificial intelligence systems, have not been transparently disclosed by the public authority (Vergiport, 2025). This prevents researchers from independently testing the absolute performance of the algorithm; accordingly, the findings presented in this article are limited to the illustrative cases and indicators made publicly available by the administration.

5. DOCUMENT ANALYSIS FINDINGS: PUBLIC INDICATORS

This section examines, within the framework of document analysis, the indicators and administrative findings shared by public authorities regarding the operational status of the KAŞIF system. Before proceeding to the analysis, it is methodologically necessary to subject the findings to a conceptual classification in order to ensure their academic validity: input, output, outcome, and impact. In this context, the system’s daily screening and analysis of 4 million taxpayers should be evaluated merely as an “operational output.” The uncovered network involving 246 million ₺ or the risky transaction volume of 276 billion ₺, by contrast, should be regarded as “short-term outcome” indicators reflecting the system’s anomaly detection capacity. However, “causal impact” measurements, such as the ultimate reduction in the shadow economy brought about by the application or the permanent structural increase it may have generated in tax compliance, fall outside the descriptive boundaries of this study and beyond the scope of currently available public indicators. Therefore, the findings are described in the context of “outcome” cases disclosed by the administration, rather than as definitive causal judgments.

5.1. Relationship-Network Findings and the Detection of Shell Companies

A prominent indicator generated by the KAŞIF system involves illustrative cases shared regarding the detection of interrelated networks of shell companies. According to the documents examined within the scope of the research, 6 of the 7 shell companies established in different tax offices under the name of a single wage-earning employee were algorithmically identified as having been created directly for the purpose of issuing fraudulent documents, and this was shared with the public (Ekonomi Gazetesi, 2026). The total volume of fraudulent invoice traffic reported to have been circulated through this network is stated as 246 million ₺ (Palabiyık & Uylaş, 2026). These data should be evaluated not as causal proof, but as publicly disclosed cases presented by the administration as examples of findings concerning relationship networks and network structures.

5.2. Time-Series and Rapid Intervention Indicators

In traditional tax auditing, the time asymmetry between the commission of an offence and its detection is a well-known structural problem in the literature. The data shared by the administration following the implementation of KAŞIF point to a notable increase in detection speed. For example, it has been reported that a taxpayer whose declared field of activity was wholesale trade, but who had no input in the form of goods purchases, was immediately marked as a “red flag” by the system after issuing three different fraudulent invoices on the first day of establishment (BirGün, 2026). Similarly, it is reported that the transactions of taxpayers who allegedly issued five fraudulent invoices amounting to 50 million ₺ as early as the fifth day following their establishment were detected in real time (Palabiyık & Uylaş, 2026).

5.3. Macro-Scale Anomaly Detection and Reported Transaction Volumes

Macro-scale indicators concerning KAŞİF's screenings on big data platforms have also been shared with the public. According to administrative bulletins, the system cross-checks tax returns, accruals, collections, foreign trade data, and Social Security Institution (SGK) data (Palabıyık & Uylaş, 2026). As a result of these screenings, 2,550 taxpayers established without any commercial, agricultural, or professional activity in the last three months were reported to be risky (Ekonomi Gazetesi, 2026). Public authorities have shared that the total risky transaction volume allegedly structured through these entities and identified by the administration amounted to 276 billion ₺ (BirGün, 2026). These figures are public indicators reflecting the screening capacity achieved by anomaly detection algorithms, rather than evidence of absolute macroeconomic prevention.

5.4. Findings Concerning Unlawful VAT Refunds

In Value Added Tax (VAT) refund processes, it is stated that capacity–volume inconsistencies of businesses requesting refunds are analyzed algorithmically. In an illustrative case reflected in the public domain, the transaction of a taxpayer who declared 7.7 million ₺ of VAT to be refunded in its tax return was reportedly blocked after the system identified inconsistencies in the taxpayer's production capacity and input–output matrices (Palabıyık & Uylaş, 2026). This statistical disclosure constitutes a descriptive finding regarding the type of detection mechanism the system operates against unlawful fund outflows detrimental to the Treasury.

5.5. Claims Concerning Accuracy Rates and Methodological Limitations

In artificial intelligence-supported algorithmic auditing mechanisms, transparency regarding false positive rates is a fundamental requirement in terms of procedural safeguards and proportionality. With respect to the performance of the KURGAN system, the pioneering component of the algorithmic architecture implemented by the Ministry of Treasury and Finance, it has been claimed/reported in the media by an expert that the model's error margin remains below 1% and that it operates with an accuracy rate of exactly 99.2% (Bloomberg HT, 2026). However, the official methodology of the machine learning algorithms used by the administration, the structure of the training datasets, and the validation protocol used to distinguish false positives are not publicly available. Therefore, rather than being treated as proven academic performance metrics in a definitive sense, such high accuracy rates should be evaluated with considerable caution and appropriate caveats as public-facing statements concerning the strength of the administrative architecture.

6. DISCUSSION

This section comparatively analyzes the structural efficiency that the KURGAN and KAŞİF algorithmic audit architectures, jointly operated by the Tax Inspection Board (VDK) and the Revenue Administration (GİB), provide for public finance in comparison with traditional rule-based systems. The ethical and socio-legal tensions generated by the quantitative and operational achievements obtained are discussed on a conceptual level, in line with Q1 journal standards, with reference to the principles of the rule of law, the presumption of innocence, and proportionality.

6.1. Traditional Auditing vs. Algorithmic Auditing Efficiency and Inter-System Synergy

In the tax auditing literature, the paradigm shift between static rule-based systems and autonomous machine learning-based structures is explained through the transformation of “time asymmetry,” or time lag, in favor of the administration. Mechanisms previously implemented in public finance, such as the Fraudulent Document Risk Analysis Program (SARP) and the Risky Refund Monitoring and Analysis Program (RİTAP), had a structurally reactive logic because they operated on the basis of predefined rules (Palabıyık & Uylaş, 2026). This reactive model was vulnerable to manipulation and functioned as a delayed structure capable of detection only months or years after public loss had already occurred.

The first and most critical evolutionary stage in the transition from this cumbersome structure to algorithmic governance began with the implementation of the Establishment-Supervised Analysis system, known as KURGAN, by the Tax Inspection Board (VDK). Playing a pioneering role in big data analytics, KURGAN is described in publicly available accounts as a proactive “early warning system” that uses massive data pools such as e-invoices, e-ledgers, and banking transactions and employs multi-criteria, relationship/network-based risk indicators among taxpayers. Thus, it initiated the artificial intelligence transformation in tax auditing (Vergiport, 2025, p. 1). This study does not reproduce this network-based mechanism; rather, it classifies the accounts presented in publicly available texts through document analysis.

Built upon the algorithmic foundation established by KURGAN, the KAŞİF system developed by GİB has carried the process into a real-time dimension through its machine learning capability. KAŞİF, which screens an average of 4 million taxpayers per day in real time across terabytes of data pools (Ekonomi Gazetesi, 2026), has produced a substantial technological synergy by integrating KURGAN's risk maps with zero-day anomaly detection. The joint operation of KURGAN's deep relational network analysis and KAŞİF's real-time anomaly detection capacity has transformed the tax administration from a traditional “watchman” role into a real-time analytical audit mechanism. Publicly available indicators and illustrative cases suggest that the selection–detection window has

narrowed and that detection speed has increased. Its causal effect on macroeconomic stability and tax compliance, however, can be tested only through comparative research designs based on anonymized administrative microdata (Palabiyik & Uylaş, 2026).

6.2. Presumption of Innocence, Administrative Sanctions, and Ethical Boundaries

Despite the operational benefits that algorithmic governance and big data analysis provide for fiscal discipline, the limited transparency of the algorithm, as noted in the methodology section, necessitates a profound ethical and socio-legal discussion in modern tax law. The operation of KURGAN/KAŞİF-type architectures as closed “black boxes” and the inability to explain the rationale for risk classification sufficiently to taxpayers may impair procedural justice in administrative processes and generate friction with the principles of legal certainty and proportionality. Therefore, “accountability and transparency” as well as “explainability and interpretability” should be treated as integral attributes of trustworthy Artificial Intelligence (AI). Meaningful transparency should be ensured, particularly at a level that enables redress in relation to erroneous or adverse outputs (NIST, 2023, pp. 2, 15–16).

Within this framework, risk scores generated by machine learning should not be positioned as decisions that, by themselves, produce sanctions. Rather, they should be treated as “preliminary indicators” or “early warnings” that prioritize audit resources and trigger verification gateways. Indeed, according to publicly available accounts of the practice, information on taxpayers “identified as being at high risk of issuing fraudulent documents” is reported to the Tax Inspection Board and simultaneously transmitted to the relevant tax offices. On-site inspections are then carried out at those taxpayers’ workplaces. Based on the decisions of commissions established at the tax offices, the registrations of individuals found to have acquired taxpayer status for the purpose of issuing fraudulent documents are subsequently closed (Palabiyik & Uylaş, 2026). This structure indicates that intermediary safeguards such as human oversight, field verification, and reasoned administrative decisions are necessary before an algorithmic score can produce legal consequences (Palabiyik & Uylaş, 2026).

Similarly, when risk analysis outputs acquire the character of a “preliminary finding,” the administration may invite the taxpayer to provide an explanation under Article 370 of the Tax Procedure Law (Vergi Usul Kanunu [VUK], 1961, Art. 370). In addition, GİB’s informational document on the invitation-to-explanation procedure summarizes procedural thresholds such as the taxpayer’s obligation to submit an explanation to the commission within 30 days after notification of the invitation letter, and the commission’s obligation to evaluate a timely explanation within no more than 45 days (Revenue Administration [GİB], 2025, p. 1). For this reason, the design of the invitation-to-explanation procedure should strengthen the “right to explanation” while avoiding a practice in which the burden of proof is effectively and disproportionately shifted onto the taxpayer because of the uncertainty inherent in algorithmic classification (VUK, 1961, Art. 370; GİB, 2025, p. 1).

The case law of the European Court of Human Rights (ECtHR) under the European Convention on Human Rights distinguishes between “ordinary tax disputes” in the field of taxation and tax sanctions that have a deterrent or punitive character, showing that Article 6 safeguards may apply to the latter category. For example, in *Bendenoun v. France*, the Court held that tax surcharges could be characterized as “criminal” in light of the combined effect of factors such as their scope and deterrent-punitive purpose, and accepted the applicability of Article 6 of the Convention (European Court of Human Rights [ECtHR], 1994, p. 16). Similarly, in *Janosevic v. Sweden*, the Court emphasized the deterrent and punitive character of the tax surcharge regime and found that it amounted to a “criminal charge” for the purposes of Article 6 (ECtHR, 2002, pp. 22–23). The ECtHR’s thematic factsheet on “Taxation” also summarizes that applications concerning taxation may engage, in particular, Article 1 of Protocol No. 1 on the protection of property and, in certain types of sanctions, Article 6 of the Convention (European Court of Human Rights [ECHR], 2023, pp. 1, 10). Within this framework, practices based solely on a “score,” lacking an explanation of reasons, and rapidly or automatically evolving into sanctions should be regarded as high-risk designs in terms of proportionality and legal certainty (ECtHR, 1994, p. 16; ECtHR, 2002, pp. 22–23; ECHR, 2023, p. 10).

Accordingly, “Explainable Artificial Intelligence (XAI)” and human oversight, including the human-in-the-loop model, should not be treated as ethical aspirations, but as institutional necessities for ensuring that algorithmic governance in tax auditing operates in accordance with the principles of the rule of law. The NIST AI RMF clearly distinguishes the “accountable and transparent” attributes of trustworthy AI from its “explainable and interpretable” attributes. Under the GOVERN function, it also positions the definition of roles and responsibilities, human–AI configurations, and oversight mechanisms through policies and procedures as a core component of governance (NIST, 2023, pp. 15–16, 21–23).

7. CONCLUSION AND RECOMMENDATIONS

This study has presented, through a holistic approach, the empirical implications of Türkiye’s transition in public finance and tax auditing processes from static rule-based structures to autonomous machine learning models such as KURGAN and KAŞİF. The maturity that algorithmic governance has reached within the framework of the Tax Administration 3.0 vision entails important policy implications and system development needs in relation to both macroeconomic stability and the boundaries of administrative law.

7.1. Policy Implications

The public indicators compiled within the scope of the study suggest that the KURGAN and KAŞİF systems have the capacity to narrow the selection–detection window and increase detection speed in tax auditing (Palabıyık & Uylaş, 2026). However, since independent validation protocols, raw microdata, and comparative evaluation designs—such as before-and-after analysis, control groups, or comparisons of similar units—are not publicly available, these figures should be interpreted with caution not as academic “performance metrics,” but as descriptive public indicators reported through detection/flagging outputs (Palabıyık & Uylaş, 2026).

This cautious approach is not merely a data-access problem specific to Türkiye. It is also a natural consequence of the fact that, as the use of AI by tax administrations rapidly expands worldwide, governance and trust dimensions are becoming increasingly decisive. Indeed, according to the OECD’s comparative data, the rate of AI use in tax administrations increased from 9% in 2016 to 69% in 2023, while an additional 24% were reported to be at the “implementation” stage (OECD, 2025a, p. 65). The OECD report on AI governance in core government functions also explicitly states that tax administrations should place particular emphasis on privacy, security, and trustworthy AI practices because of their capacity to process sensitive data (OECD, 2025b, p. 170). The report further emphasizes that taxpayer trust is a critical element in the functioning of tax administration and that the AI used in this context must therefore be operated in a trustworthy, transparent, and accountable manner (OECD, 2025b, p. 176). Accordingly, at the policy level, the goal of “greater analytical capacity” should be considered together with the simultaneous goal of “stronger procedural safeguards and governance architecture.”

Within this framework, the capacity for autonomous anomaly detection and risk prioritization may be considered capable of generating a deterrent effect in terms of public finance policies. Nevertheless, the direction and magnitude of the causal effect on deterrence and tax compliance cannot be determined conclusively without empirical testing because of the “power–trust” interaction emphasized by the Slippery Slope Framework. An increase in audit power may suppress risky behavior in the short term; however, if decision-making processes remain “black boxes” and explainability and contestability remain weak, this may erode the trust channel and adversely affect voluntary compliance in the medium and long term (Kirchler, Hoelzl, & Wahl, 2008, pp. 211–212).

Moreover, the modern tax auditing literature particularly emphasizes selection bias—selective labels/endogenous selection—as one of the main methodological problems encountered in evaluating machine learning-based audit systems. Since “ground truth” labels are often observable only for audited taxpayers, it becomes difficult to reliably estimate general compliance effects and model performance without independent data or comparative research designs (Battaglini et al., 2022, pp. 1–3). Therefore, descriptive frameworks based on public indicators, such as the present study, serve not to prove “impact,” but rather to make visible the institutional safeguard needs generated by system design.

Accordingly, the most rational policy-level inference is that the “early warning/risk prioritization” function of these systems should be strengthened while procedural safeguards—explainability, objection/review mechanisms, false-positive management, human oversight, and audit trails—are institutionalized simultaneously. In this context, the principles of the human-in-the-loop model and Explainable Artificial Intelligence (XAI) should no longer be treated as abstract ethical aspirations. The IMF technical note prepared for tax and customs administrations provides a practical roadmap for responsible AI use by placing key control points at the core of governance, including the formulation of an institutional AI policy, maintenance of an AI use case inventory, legal and ethical review, preservation of HITL/HOTL mechanisms where appropriate, risk assessment before deploying new use cases, and publication of the inventory for transparency purposes (Aslett et al., 2024, pp. 20–22). In the Turkish context, a hybrid administrative safeguard mechanism must also be implemented and enforced without compromise: algorithmic outputs, or black-box scores, should not directly and autonomously turn into administrative penalties; findings must be verified through concrete audit procedures such as on-site inspections and expert tax-inspector investigations; and final authority must remain with the human auditor.

7.2. Future Research Recommendations and System Development

The extensive audit power that algorithmic governance and big data analytics provide to the tax administration is subject to ethical and transparency concerns arising from the operation of these systems as closed “black boxes.” Future academic studies and system development policies should focus on understanding the nature of false positive cases produced by machine learning models. In this context, in order for algorithmic biases and error margins to be subjected to scientific testing by independent researchers, anonymized open datasets should be made available to the academic community through technical methods that strictly do not violate the principle of “tax confidentiality.”

On the other hand, the direct transformation of automatic decisions based solely on data mining and algorithmic risk scoring into severe administrative sanctions, such as the ex officio termination of taxpayer registration, carries the risk of potential grievances in relation to the presumption of innocence and the principle of proportionality. To minimize this socio-legal tension, artificial intelligence applications should be positioned not as fully autonomous decision-making authorities, but as early warning and risk prioritization tools. In high-risk cases, a hybrid model in which the human-in-the-loop/human oversight approach is mandatory for

ensuring procedural justice, and in which the final decision is made exclusively through the discretion and legal reasoning of expert tax auditors, is an indispensable legal requirement for the sustainability of the system (Özmumcu, 2026).

DECLARATIONS

Ethics Approval and Consent to Participate

Not applicable. This study is based exclusively on publicly available documents, official reports, institutional publications, and publicly accessible statistical data. No human participants, personal data, or confidential records were involved.

Availability of Data and Materials

All data analyzed during this study are derived from publicly available official reports, institutional publications, and other publicly accessible sources cited in the reference list. No proprietary datasets were used.

Competing Interests

The author declares that there are no competing interests.

Funding

The author received no financial support for the research, authorship, or publication of this article.

Authors' Contributions

S.A. conceived the study, conducted the literature review, collected and analyzed the data, interpreted the findings, and wrote and approved the final manuscript.

Acknowledgements

The author would like to thank the institutions whose publicly available reports and statistical publications contributed to this study.

Conference Presentation

An earlier version of this study (abstract only) was presented at the 13th International CEO Social Sciences Congress, held at the New Delhi Institute of Management, India, on 6–7 June 2026. Participants had the opportunity to attend the congress online. The present article is a substantially expanded and revised version prepared for journal publication.

Publisher's Note

The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Aslanpınar, B. (2026). *KURGAN nedir? | Maliye'den yazı geldiyse ne yapmalı? | Vergi cezasına karşı pişmanlık mı itiraz mı?* [What is KURGAN? | What should be done if a letter from the Ministry of Finance has arrived? | Remorse or objection against a tax penalty?] [Video]. YouTube. <https://www.youtube.com/watch?v=GtOEcrYxc0Q>
- Aslett, J., González, I., Hadwick, D., Hamilton, S., Hardy, M., & Pérez, A. (2024). *Understanding artificial intelligence in tax and customs administration* (Technical Notes and Manuals No. 2024/06). International Monetary Fund. <https://www.imf.org/-/media/files/publications/tnm/2024/english/tnmea2024006.pdf>
- Battaglini, M., Guiso, L., Lacava, C., Miller, D. L., & Patacchini, E. (2022). *Refining public policies with machine learning: The case of tax auditing* (NBER Working Paper No. 30777). National Bureau of Economic Research. <https://doi.org/10.3386/w30777>
- BirGün. (2026, January 20). *Bakan Şimşek duyurdu: Vergide yeni dönem* [Minister Şimşek announced: A new era in taxation]. <https://www.birgun.net/haber/bakan-simsek-duyurdu-vergide-yeni-donem-685783>
- Bloomberg HT. (2026, January 26). *Sahte fatura sorununda sona gelindi* [The fraudulent invoice problem is nearing its end]. <https://www.bloomberght.com/sahte-fatura-sorununda-sona-gelindi-3767384>
- Çember, E. (2026). *Sahte faturaya dijital fren: KAŞIF* [A digital brake on fraudulent invoices: KAŞIF]. Çember Yeminli Mali Müşavirlik. <https://cemberymm.com.tr/Sahte-Faturaya-Dijital-Fren-KASIF-748u.html>

7. Dalglish, S. L., Khalid, H., & McMahon, S. A. (2020). Document analysis in health policy research: The READ approach. *Health Policy and Planning*, 35(10), 1424–1431. <https://doi.org/10.1093/heapol/czaa064>
8. Ekonomi Gazetesi. (2026, January 20). “KAŞİF” ile günde 4 milyon mükellef inceleniyor [Four million taxpayers are examined daily with “KAŞİF”]. <https://www.ekonomigazetesi.com/ekonomi/kasif-ile-gunde-4-milyon-mukellef-inceleniyor-70337>
9. European Court of Human Rights. (1994, February 24). *Case of Bendenoun v. France* (Application No. 12547/86). <https://hudoc.echr.coe.int/app/conversion/pdf/?filename=001-57863.pdf&id=001-57863&library=ECHR>
10. European Court of Human Rights. (2002, July 23). *Case of Janosevic v. Sweden* (Application No. 34619/97). <https://hudoc.echr.coe.int/app/conversion/pdf/?filename=001-60628.pdf&id=001-60628&library=ECHR>
11. European Court of Human Rights. (2023, March). *Factsheet – Taxation and the European Convention on Human Rights*. Council of Europe. https://www.echr.coe.int/documents/d/echr/fs_taxation_eng
12. Gelir İdaresi Başkanlığı. (2025). *İzaha davet uygulaması* [The invitation-to-explanation procedure] [Infographic]. T.C. Hazine ve Maliye Bakanlığı. https://cdn.gib.gov.tr/api/gibportal-file/file/getFileResources?objectKey=arsiv%2Fyardim-kaynaklar%2Finfografikler%2Fpdfs%2F2025_IzahaDavet.pdf
13. Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The “slippery slope” framework. *Journal of Economic Psychology*, 29(2), 210–225. <https://doi.org/10.1016/j.joep.2007.05.004>
14. National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
15. OECD. (2016). *Advanced analytics for better tax administration: Putting data to work*. OECD Publishing. <https://doi.org/10.1787/9789264256453-en>
16. OECD. (2017). *The changing tax compliance environment and the role of audit*. OECD Publishing. <https://doi.org/10.1787/9789264282186-en>
17. OECD. (2020). *Tax administration 3.0: The digital transformation of tax administration*. OECD Publishing. <https://doi.org/10.1787/ca274cc5-en>
18. OECD. (2025a). *Tax administration 2025: Comparative information on OECD and other advanced and emerging economies*. OECD Publishing. <https://doi.org/10.1787/cc015ce8-en>
19. OECD. (2025b). *Governing with artificial intelligence: The state of play and way forward in core government functions*. OECD Publishing. <https://doi.org/10.1787/795de142-en>
20. Özmumcu, M. O. (2026). *Vergilemede kaygan zemin teorisi ve enflasyonla ilişkisi*. [The slippery slope framework in taxation and its relationship with inflation]. Mucteba Onurhan Özmumcu. <https://www.muctebaonurhanozmumcu.com/vergilemede-kaygan-zemin-teorisi-ve-enflasyonla-iliskisi>
21. Palabıyık, D. Ç., & Uylaş, H. Ö. (2026, January 20). *Naylon fatura, yapay zekâ tabanlı KAŞİF ile anında yakalanıyor* [Fraudulent invoices are detected instantly with AI-based KAŞİF]. Anadolu Ajansı. <https://www.aa.com.tr/tr/ekonomi/naylon-fatura-yapay-zeka-tabanlı-kasif-ile-anında-yakalanıyor/3804685>
22. Softa, M. (2025, October 24). *Vergi incelemelerinde yeni yaklaşımlar: Yapay zekâ denetimleri ve KURGAN uygulamaları* [New approaches in tax inspections: Artificial intelligence audits and KURGAN applications]. Vergiport. <https://vergiport.com/blog/vergi-incelemelerinde-yeni-yaklasimlar-yapay-zeka-denetimleri-ve-kurgan-uygulamaları>
23. Uyumsoft. (2026, January 27). *Vergi denetiminde yapay zekâ destekli yeni takip ve analiz sistemi: KAŞİF* [A new AI-supported monitoring and analysis system in tax auditing: KAŞİF]. <https://www.uyumsoft.com/blog/vergi-denetiminde-yapay-zeka-destekli-yeni-takip-ve-analiz-sistemi-kasif>
24. Vergi Usul Kanunu. (1961). *Kanun No. 213* [Tax Procedure Law, Law No. 213]. <https://hukukmusavirligi.diyaret.gov.tr/Documents/213%20Say%C4%B1%C4%B1%20Vergi%20Usul%20Kanunu.pdf>
25. Yıldız, F., & Köşer, Ö. (2025). Yapay zekâ uygulamalarının vergi gelirlerine etkisi [The effect of artificial intelligence applications on tax revenues]. *Vergi Raporu*, 313, 7–23. <https://vergiraporu.com.tr/upImage/org/1f0c9b0.PDF>