



ARTICLE



<https://doi.org/10.1057/s41599-025-06099-7>

OPEN

# Balancing innovation and integrity: AI in tax administration and taxpayer rights

Evgeny Guglyuvatyy<sup>1,2</sup>✉

Artificial intelligence (AI) is transforming tax administration by improving efficiency, compliance, and decision-making. However, this shift raises critical concerns about transparency, accountability, and taxpayer rights. This paper examines how AI-driven systems impact legal fairness, due process, and the integrity of tax procedures. It highlights risks such as algorithmic bias, opacity, and weakened procedural safeguards, while acknowledging AI's potential to streamline enforcement. To safeguard taxpayer rights, the paper proposes an independent AI oversight mechanism to explain and review tax decisions. This system would enhance transparency, reinforce trust, and ensure legal accountability. The paper calls for regulatory frameworks that embed oversight, uphold public trust, and balance innovation with fundamental legal protections.

<sup>1</sup> Monash University, Subang Jaya, Malaysia. <sup>2</sup> Faculty of Law, University of Technology Sydney, Sydney, Australia. ✉email: [evgeny.guglyuvatyy@monash.edu](mailto:evgeny.guglyuvatyy@monash.edu)

## Introduction

Artificial Intelligence (AI) has become increasingly pervasive across diverse sectors, leveraging expansive digital repositories and sophisticated algorithms to process complex queries, generate conclusions, and simulate human-like interactions. In tax administration, AI progressively automates decisions, analyses datasets, and enhances compliance monitoring (Aslett et al. 2024). The evolution of AI in taxation began several decades ago with foundational systems like ‘Taxman’ (McCarty 1977). Since then, applications have expanded to include taxpayer education, fraud detection, and compliance monitoring. By the late 1990s, machine visioning technology facilitated automated customs inspections, while the last three decades have seen widespread adoption of advanced analytics and machine learning (ML) techniques due to expanding electronic tax data (Bruckner and Coil 2024; Witten et al. 2016).

A significant recent development is Generative AI (GenAI), which enables the creation of human-like content, including text, images, and videos. The integration of GenAI into tax administration coincides with broader digital transformation initiatives, reshaping compliance dynamics for both taxpayers and authorities (Brynjolfsson and McAfee 2017). The implications of automated systems, particularly generative AI, in the field of taxation are multifaceted and complex. While it can provide individual taxpayers with real-time advisory support to improve compliance and reduce reliance on costly professional services, these same capabilities pose risks when exploited for illicit purposes. AI-driven tools can optimise tax fraud schemes by providing guidance on jurisdictions, business structures, and legal frameworks with minimal detection risk (Zheng et al. 2024). AI-assisted fraud may further evade scrutiny through strategic document modification based on AI-generated insights (Zheng et al. 2024).

As AI transforms tax administration, it presents both opportunities and challenges. AI-driven automation, predictive analytics, and natural language processing enhance administrative efficiency and revenue collection but raise concerns regarding data security, legal accountability, and algorithmic bias. Policies balancing innovation with robust governance mechanisms are essential to harness AI’s benefits while mitigating risks (OECD 2025).

Within this context, the protection of taxpayer rights emerges as a central issue. Taxpayer rights generally refer to the legal guarantees ensuring fairness, transparency, and due process in the relationship between taxpayers and the state. Importantly, these rights also encompass the ability of taxpayers to engage in legitimate tax planning and to select their economic circumstances in ways that minimise tax liability. It is well established that taxpayers are not obliged to structure their affairs in a manner that maximises the tax burden; rather, they may legitimately choose arrangements that reduce it (Kollruss 2023; Kollruss 2025a). Thus, safeguarding taxpayer rights requires recognising both procedural guarantees and substantive entitlements, including the right to organise one’s economic activities lawfully to mitigate taxation.

This paper examines AI’s role in tax administration and its implications for taxpayer rights. It analyses approaches to AI adoption, identifies key challenges and risks associated with automation in tax enforcement, and evaluates the extent to which AI-driven tax administration aligns with principles of fairness, transparency, and due process. Finally, the paper provides recommendations aimed at balancing technological innovation with taxpayer protections to ensure AI enhances rather than undermines tax compliance and legal integrity.

## AI and tax administration

The rapid integration of artificial intelligence has prompted regulatory initiatives at both national and international levels to address legal and ethical concerns while enhancing efficiency and accuracy. As outlined in Table 1, regulatory frameworks have emerged across jurisdictions, with stakeholders seeking to provide legal certainty and governance structures for AI applications.

While AI presents potential risks to tax systems, it also offers opportunities to enhance administration by improving fairness, collection efficiency, and anti-evasion measures. A significant aspect of AI’s role is optimising enforcement, improving taxpayer services, and enhancing compliance. In the United States, the Inflation Reduction Act allocated substantial resources to modernise the Internal Revenue Service (IRS) through AI-driven tools, facilitating more efficient return processing, anomaly detection, and streamlined enforcement (Werfel 2023). IRS Commissioner Daniel Werfel credited these advancements with improving the taxpayer experience during the 2023 filing season. Similarly, the OECD has explored AI’s role in tax administration as part of broader digital economy reforms (OECD 2025).

Despite these benefits, AI-driven tax administration presents significant legal and ethical challenges. Ensuring transparency is crucial for maintaining public trust and preventing procedural injustices (Amparo and Ruiz 2022). A key concern is the risk of biased or opaque decision-making, as illustrated by the landmark ruling by the District Court of The Hague on February 5, 2020, concerning the Dutch System Risk Indication (SyRI) system. The court found that SyRI violated the right to privacy under Article 8(2) of the European Convention on Human Rights, citing a lack of transparency in how risk profiles were generated (Nederlands Juristen Comite Voor de Mensenrechten 2020). The ruling reinforced the principle that AI-driven tax enforcement must adhere to fundamental legal rights, including non-discrimination and due process.

The ruling was further catalysed by revelations that a tax assessment algorithm had relied on nationality-based selection criteria without a lawful basis, leading to discriminatory outcomes. The fallout prompted government intervention and drew condemnation from the UN Special Rapporteur on extreme poverty and human rights, who warned that AI could amplify existing social inequalities and institutional biases

**Table 1 Key AI regulatory initiatives.**

Jurisdiction	Initiative	Scope
European Union (EU)	AI Act Draft Proposal (Jan 21, 2024)	AI systems, including GenAI
Council of Europe	Draft Framework Convention on AI (Dec 18, 2023)	AI systems, including GenAI
Colombia	Proyecto de Ley Estatutaria 059 (Nov 14, 2023)	AI systems
OECD	AI Principles (Nov 8, 2023)	AI systems, including GenAI
United States	Executive Order on AI (Oct 30, 2023)	AI systems, including dual-use applications
China	Interim Measures on GenAI (Aug 15, 2023)	Public use of GenAI applications
South Korea	AI Responsibility and Regulation Bill 2123709 (Aug 8, 2023)	AI systems, including GenAI

(Alston et al. 2019). This case exemplifies the dangers of unchecked algorithmic governance and highlights the urgent need for transparency, accountability, and legal oversight in AI-driven public administration.

While SyRI is not a generative AI system, it remains a powerful example of automated decision-making technology with significant relevance to the current discourse on GenAI in taxation. Given the novelty of GenAI and the scarcity of public case studies involving its use by tax authorities, earlier systems like SyRI offer critical insight into the potential risks, legal pitfalls, and governance challenges that may similarly arise with GenAI's deployment.

AI-driven decision-making undoubtedly offers opportunities to improve the efficiency and accuracy of tax administration. However, integrating AI into automated decision-making (ADM) processes also introduces significant regulatory and ethical challenges. In many jurisdictions, current legal frameworks are not adequately equipped to manage the complexities of AI-based ADM, especially in contexts requiring nuanced, discretionary judgements (Amparo and Ruiz 2022). As AI assumes a greater role in tax determinations and enforcement, maintaining transparency, accountability, and procedural fairness becomes increasingly critical.

In Australia, the integration of AI into tax administration is being driven by technological advancements, expanded data access, evolving tax policies, and shifting societal expectations. As of May 2024, the Australian Taxation Office (ATO) had 43 AI models in production, alongside eight generative AI tools approved for public use (ATO 2024a). These AI applications support tax assessments, anomaly detection, and compliance monitoring, processing approximately 36 million documents to identify entities of interest (ATO 2024b). To ensure secure implementation, the ATO has established governance mechanisms, including mandatory staff instructions, data ethics principles, privacy impact assessments, and security risk evaluations.

Despite these measures, significant gaps persist. According to an Australian National Audit Office report, mechanisms for assessing the effectiveness of the ATO's strategy remain underdeveloped, and 74% of AI models in production lack completed data ethics assessments (Australian National Audit Office 2024). Even more concerning, while human oversight remains formally in place for adverse tax decisions, the ATO has acknowledged a declining ability among staff to explain AI-driven outcomes, highlighting the urgent need to preserve meaningful human intervention and accountability.

Globally, tax administrations operate through function-based, taxpayer segment-based, or tax-type-based models (Cumberland 2024), within which AI is increasingly deployed to enhance efficiency, detect fraud, and automate service delivery. Tools like chatbots, automated risk profiling, and predictive analytics are now integral to optimising workflows and identifying non-compliance (Cumberland 2024). Despite these advancements, AI has not yet been extensively applied to administrative decision-making within tax authorities. However, its broader deployment is not a matter of if, but when. AI offers unparalleled speed and efficiency in processing information and allocating resources, making its integration into decision-making processes both logical and inevitable (Nathwani 2024).

Yet, this transition poses significant structural and legal challenges. The current tax administrative framework is ill-equipped to accommodate AI in discretionary decision-making roles. Moving from human-led interpretation to algorithmic decision-making would fundamentally alter the basis of tax authority and discretion, marking not just a technical upgrade but a legal and institutional transformation (Nathwani 2024).

A critical distinction must be made between AI serving as a decision-support tool and AI autonomously determining taxpayer obligations. As tax administrations increasingly rely on automated systems, concerns arise regarding the potential erosion of legal expertise in decision-making. Without appropriate safeguards, tax authorities risk shifting interpretative power from legally trained professionals to technology developers who may lack the necessary legal acumen. Some jurisdictions, such as France, are exploring initiatives to codify tax rules into programming languages while ensuring AI models adhere to established legal norms (Mohun and Roberts 2020).

The opacity of AI decision-making further challenges taxpayer rights. Taxpayers must be able to understand and contest AI-generated determinations, particularly when these affect fundamental rights such as privacy and access to fair administrative processes. AI's reliance on complex, proprietary algorithms can obscure the rationale behind tax assessments, making it difficult for individuals to challenge erroneous or biased decisions. This necessitates robust procedural safeguards, including mechanisms for explainability and independent oversight (Bardutzky 2022).

Explanations in AI-driven tax administration serve three key stakeholders: tax authorities ensuring compliance with legal and ethical standards, taxpayers requiring transparency to understand and contest decisions, and oversight bodies safeguarding procedural fairness (Fink and Finck 2022). AI has the potential to revolutionise tax administration by enhancing efficiency, reducing evasion, and improving compliance. However, its deployment must be carefully managed to uphold legal, ethical, and governance standards. Ensuring AI-driven decision-making aligns with principles of fairness, transparency, and accountability requires robust regulatory frameworks and continuous oversight mechanisms.

### AI tax administration and taxpayers' rights

Artificial intelligence is rapidly transforming tax administration, but legal safeguards and accountability mechanisms are not keeping pace. The EU AI Act, while drawing a distinction between General Purpose AI and high-risk AI systems, implicitly acknowledges that even ostensibly low-risk AI can produce harmful outcomes when deployed in sensitive domains (EU AI Act). Crucially, the Act imposes transparency obligations on systems that interact with individuals or generate content capable of misleading or deceiving users, regardless of their formal classification (European Parliament 2025). This is an important step reflecting the EU's proactive stance on preventing harm even outside traditionally defined high-risk scenarios (European Parliament 2025). Yet, when it comes to tax administration, a critical gap remains.

The legal and institutional frameworks governing tax authorities are poorly equipped for the integration of AI into discretionary decision-making processes. As Nathwani (2024) argues, the application of AI in such contexts does not merely optimise existing systems; it reconfigures the very nature of public authority, blurring lines of responsibility, undermining accountability, and posing serious risks to procedural fairness and due process. Despite the escalating use of AI tools in tax enforcement, most jurisdictions lack targeted legal safeguards to protect taxpayer rights. Even the EU AI Act, with its ambitious regulatory vision, does not directly address AI's role in tax administration, leaving a regulatory vacuum at the intersection of automated governance and individual rights.

The effective integration of AI into tax administration necessitates a comprehensive and balanced approach. This balance requires robust data governance, transparent legal frameworks, and multi-stakeholder collaboration to ensure AI-driven systems

align with societal values and ethical principles. Given increasing reliance on AI in tax enforcement, regulatory oversight must address concerns related to algorithmic transparency, data privacy, and accountability of automated decision-making processes (Saba and Monkam 2024).

A fundamental concern is the extent to which automated decision-making processes remain transparent and subject to scrutiny, and is in line with the right of the taxpayer to be informed. Scholars have speculated that future advancements may enable government agencies to formulate legal norms autonomously (Coglianese and Lehr 2017). This possibility is not unfounded, as tax authorities increasingly employ AI to automate data analysis, indirectly shaping legal interpretations through technological applications. However, entrusting extensive datasets to AI-driven systems necessitates stringent safeguards to prevent potential infringements on taxpayer rights.

While many jurisdictions have formal instruments like taxpayers' charters or bills of rights, these protections are often non-binding and inconsistently enforced. The IBDF (2020) found that although 58% of surveyed jurisdictions have such instruments, only 47% make them legally binding, leaving the right to transparency and explanation vulnerable to administrative erosion (IBDF 2020). One of the core components of these instruments is the right to be informed. For instance, U.S. taxpayers are entitled to clear explanations of tax laws, administrative procedures, and IRS decisions (*IRS Taxpayer Bill of Rights*). Similarly, in Australia, the Australian Taxation Office is required to communicate clearly across all official documents and interactions, including forms, notices, and correspondence (ATO Our Charter n.d.).

This right to be informed is closely tied to public trust, which is essential to effective tax administration. Public confidence significantly enhances voluntary compliance, particularly in the context of emerging technologies (Bentley 2022). Research demonstrates that trust in tax authorities significantly influences taxpayer behaviour, particularly with emerging technologies (Wachter et al. 2017). The deployment of AI raises significant concerns regarding transparency, fairness, and accountability, particularly when decisions impact taxpayers' rights and obligations. There remains ongoing debate about the extent to which there is a genuine right to an explanation for affected individuals.

Transparency plays a crucial role in addressing these concerns. The so-called 'right to explanation' is intended to ensure that taxpayers can understand and evaluate the rationale behind AI-generated assessments. This right is integral to procedural fairness, enabling taxpayers to contest AI-driven determinations (Bal 2019). Without transparency, the opacity of AI systems may hinder due process, particularly when predictive analytics influence tax liabilities and penalties. Indeed, the legitimacy of using predictive analytics in determining tax obligations has already been questioned in several jurisdictions, highlighting the need for clearer legal and ethical boundaries.

Beyond transparency, fairness is critical in AI-driven tax administration (Ng et al. 2020). Non-discrimination and equal treatment are foundational principles of tax law, requiring authorities to ensure that automated decisions do not result in arbitrary or biased outcomes. Yet, AI systems trained on imbalanced or unrepresentative datasets may inadvertently reproduce or even amplify existing societal biases. If an AI system is trained on data overrepresenting specific demographics, it may perpetuate discriminatory patterns. Addressing these biases is imperative to upholding the integrity of AI-driven tax assessments.

Fairness also intersects with broader taxpayer rights, including procedural fairness, the right to an explanation, data protection, and access to remedies (Gorski et al. 2024). Moreover, taxpayer rights extend beyond statutory guarantees to include the right to

invoke higher-ranking sources of law. As Kollruss (2025b) notes, taxpayers may rely on supranational law, such as European Union law, as well as domestic constitutional law, when challenging decisions of tax authorities. National law, and tax law in particular, must comply with or be consistent with these higher-ranking legal frameworks (Kollruss 2025c). This principle ensures that automated tax enforcement remains subject not only to administrative safeguards but also to constitutional and supra-national oversight, thereby reinforcing the rule of law. These safeguards become especially critical when algorithmic opacity limits a taxpayer's ability to understand or challenge automated decisions (Zalnierute et al. 2021). Ensuring transparency in decision-making is not just an ethical imperative - it is necessary for maintaining the rule of law in tax administration.

Legal and institutional accountability mechanisms must evolve accordingly. Under the EU AI Act, tax administrations may be classified either as AI providers, tasked with ensuring data quality, risk management, and conformity assessments (Articles 16–29), or as AI deployers, responsible for human oversight, performance monitoring, and transparency (Articles 29–31). When administrations perform both roles, they must adhere to the full spectrum of regulatory obligations across the AI lifecycle (Veale and Borgesius 2021). The OECD's Framework for the Classification of AI Systems reinforces this role-based distinction, advocating for risk-sensitive and context-specific regulation in high-stakes domains such as taxation (OECD 2022).

Accountability, therefore, must be both structural and operational (Gorski et al. 2024). In the public sector, accountability means that a designated authority bears responsibility for AI-generated outcomes. Taxpayers must have accessible and effective avenues for redress when automated systems make errors. Human review of AI-generated tax decisions is a critical safeguard, but its effectiveness hinges on the capacity of tax officials to meaningfully interpret and override complex algorithmic outputs. The technical opacity of advanced AI models can limit this capacity, underscoring the need for clearly defined responsibilities, transparent review processes, and robust oversight mechanisms.

These concerns are not limited to tax administration. The broader shift toward a 'digital welfare state', where governments increasingly rely on algorithmic tools to deliver public services, raises similar challenges. A UN report (2019) warns that such models may entrench bias and erode accountability, especially in the absence of strong regulatory safeguards. While AI is often perceived as neutral or objective, its outputs can mirror the biases encoded in training data or algorithm design. Without comprehensive legal frameworks that prioritise human rights and due process, automated decision-making risks reinforcing inequality. In the tax context, this places a proactive duty on administrations to embed fairness and accountability into system design and implementation, even in the absence of explicit statutory mandates.

Another critical challenge is striking a balance between algorithmic transparency and protecting proprietary intellectual property (Nathwani 2024). The requirement to disclose AI models' inner workings presents legal and ethical dilemmas, particularly when tax authorities rely on external technology providers. This issue is complicated when AI service providers operate across multiple jurisdictions, necessitating international cooperation to establish standardised transparency norms.

AI-based taxpayer profiling raises additional concerns. Tax administrations increasingly employ AI to categorise taxpayers based on variables such as past filings, business activities, and geographic locations. While such classifications enhance enforcement efficiency, they also risk stigmatisation and discriminatory treatment. Individuals or businesses flagged as 'high-risk' may

face disproportionate scrutiny, negatively impacting their credit ratings and financial standing. The reliability of AI-driven classifications depends on training data integrity and statistical methodology validity. Consequently, tax authorities must ensure rigorous testing to minimise biases and prevent unfair treatment.

The risks of AI-driven tax administration are exemplified by two Australian cases: Robodebt (The Royal Commission into the Robodebt Scheme 2023) and Pintarich (Pintarich v Deputy Commissioner of Taxation 2018). The Robodebt scheme involved an automated compliance system used by Centrelink and the Australian Taxation Office to identify welfare overpayments. The system relied on a data-matching algorithm that averaged ATO records of fortnightly earnings, often leading to erroneous debt assessments. This flawed approach ignored income fluctuations and improperly shifted the burden of proof onto individuals, violating established legal norms. The scheme was ultimately ruled unlawful, highlighting the perils of automated decision-making without adequate safeguards.

Similarly, Pintarich underscores the potential for AI-driven systems to generate legally ambiguous outcomes. The ATO's automated letter-generation system issued a notice indicating a specific tax outcome, which an officer later contended was not an actual decision. The Full Federal Court ruled that no binding decision had been made, warning of the dangers of treating machine-generated communications as autonomous decisions without human oversight. This case emphasises the necessity of clear legal responsibility frameworks in AI-assisted decision-making.

The Australian Human Rights Commission has established key principles governing AI-driven decision-making, emphasising compliance with international human rights standards, harm minimisation, and implementation accountability (Australian Human Rights Commission 2021). The Commission stresses that AI-based decision-making should be subject to robust regulation to prevent adverse consequences.

As AI-driven decision-making evolves, legal and ethical frameworks must adapt to ensure fairness, transparency, and accountability. The principles of administrative law must be rigorously upheld to prevent AI from undermining procedural fairness and eroding public trust. The broader implications of AI integration necessitate an approach that preserves foundational principles of tax governance. The principle of 'no taxation without representation' must not be supplanted by 'no taxation without automation' (Denver University Law Review 2011). Instead, AI should function as an adjunct to, rather than a replacement for, established principles of tax justice and procedural fairness.

### Future prospects and recommendations

The future of AI-driven tax administration depends on balancing innovation with fundamental taxpayer rights, particularly regarding transparency, accountability, and procedural fairness. A critical aspect is ensuring AI systems provide clear and interpretable justifications for determinations. Transparency is not merely technical but a legal and ethical imperative, as taxpayers must be able to understand and challenge AI-driven tax decisions. Without adequate transparency and explainability, individuals may be unable to dispute assessments, ultimately undermining due process and eroding trust in tax institutions.

Recent regulatory developments support this imperative. Under Article 86 of the EU AI Act, individuals impacted by high-risk AI systems are entitled to meaningful information about how those systems function, a right rooted in long-standing administrative law principles of transparency and accountability (Wachter et al. 2017; Edwards and Veale 2017). In tax

administration, this includes explaining how the system identifies risks (such as fraud), what conclusions it reaches, and the extent to which those outputs influence final decisions. These disclosures are especially vital when serious allegations, like tax evasion, are made. They empower taxpayers to contest flawed or opaque outcomes and seek meaningful human review (Zalnieriute et al. 2021). This right should be enforceable through existing access-to-information mechanisms and appeals processes.

Complementing this, Article 65 of the proposed AI Liability Directive offers individuals the right to redress for harms caused by faulty AI outputs, such as false positives in fraud detection systems (Veale and Borgesius 2021). To ensure these rights are effective, legal standards must define what constitutes an adequate explanation, when it should be provided, and how affected parties can seek remedies. Such standards must be coupled with practical and accessible pathways for remedy.

Despite these emerging protections, significant implementation gaps remain. Many existing tax lodgement systems are not user-friendly and often require professional assistance even for routine filings. In numerous jurisdictions, small businesses and individuals face structural barriers to challenging tax authority decisions, compounded by the opacity of automated determinations (Myburg and Schutte 2025). Moreover, most tax administrations still lack dedicated policies or frameworks governing the design, development, and deployment of AI tools. Instead, they rely on generic enterprise policies that do not adequately address the distinct risks posed by automated decision-making in taxation.

This governance deficit is further exacerbated by the insufficient integration of ethical and legal principles into AI development pipelines. Key safeguards such as fairness, bias mitigation, reliability, privacy protection, transparency, contestability, and accountability are often inconsistently applied or entirely absent. In addition, many administrations have yet to establish structured assurance mechanisms, including processes for testing, validating, reviewing, and documenting decisions made throughout the AI model development.

Although efforts are underway to enhance the transparency, fairness, and accountability of AI in tax administration, the pace and effectiveness of these reforms remain uncertain. The rapid deployment of AI systems often outpaces the development of appropriate regulatory and oversight frameworks. This lag creates an urgent need for robust systems that not only explain how decisions are made but also allow for rigorous scrutiny of those decisions with respect to fairness, accuracy, and legal compliance.

At the same time, existing mechanisms for ensuring equitable access to tax justice, such as traditional tax and law clinics, remain limited in scope. These clinics typically serve a narrow segment of the population, constrained by geographic, financial, and institutional barriers (Afield 2019). As a result, many low-income taxpayers face significant obstacles in accessing timely, affordable, and accurate tax advice. The integration of independent AI-driven systems offers the potential to transform this access gap. By interfacing with tax authority databases and existing legal frameworks, such technologies could provide scalable, real-time, and multilingual guidance tailored to diverse taxpayer needs. These systems have the potential to democratise access to tax assistance and improve procedural fairness, particularly for underserved populations.

In practice, some jurisdictions are already exploring this potential. The United States Internal Revenue Service has experimented with AI to enhance compliance and taxpayer services, though its full potential remains untapped in democratising access (Internal Revenue Service 2022). Similarly, the European Union's "Fiscalis Taxation Programme" encourages member states to adopt digital tools supporting taxpayers, with projects like Estonia's AI-based tax chatbot 'Kratt' demonstrating the

feasibility of national-scale, user-oriented systems (European Commission 2020).

Properly implemented, such AI-based platforms could serve as virtual “first stops” for taxpayers with limited resources or digital literacy by delivering plain-language explanations of tax rules, filing obligations, and entitlements. This promotes access to justice and procedural fairness while aligning with taxpayer rights principles articulated by bodies like the Taxpayer Advocate Service and European Taxpayers’ Code. An AI system, if transparently governed and regularly audited, can offer a reliable alternative to costly private advisory services or under-resourced public clinics.

To facilitate fairness and accountability, there is a compelling need to develop an independent, AI-powered oversight mechanism for reviewing automated tax determinations, assessments, and administrative decisions. Such a system should function as an external auditing tool, cross-verifying logic, data inputs, and legal interpretations in AI-driven tax decisions. Unlike conventional validation processes, this oversight mechanism must evaluate whether generated determinations align with broader legal, ethical, and procedural safeguards. Specifically, the tool should:

1. Provide transparent explanations offering detailed, comprehensible insights into how tax determinations were reached, referencing statutory law, judicial precedents, and administrative practices.
2. Identify and correct algorithmic bias by assessing historical data patterns to detect and mitigate biases leading to discriminatory assessments.
3. Ensure procedural fairness and due process by facilitating taxpayer engagement, making automated decisions contestable, and ensuring legal reasoning is documented for appeals.

The need for such oversight is underscored by well-documented risks of algorithmic bias in AI models. AI-driven tax systems trained on historical enforcement data may perpetuate past biases, leading to unfair outcomes that disproportionately impact vulnerable taxpayers (Barocas et al. 2019). An independent AI-based review mechanism that continuously monitors and corrects biases is essential for equitable tax administration. Such a system can provide explanatory mechanisms for taxpayer decisions and enhance procedural justice in tax assessments and disputes, particularly for those lacking resources for professional tax agents.

Accountability remains central to effective tax administration. Historically, tax decision-making has involved responsible human authorities to whom taxpayers can appeal disputed assessments. However, as AI assumes more prominence in determinations, accountability faces new challenges. ‘Automation bias,’ where officials overly rely on AI-driven decisions without sufficient scrutiny, poses a significant risk (Binns and Veale 2021). In complex matters requiring subjective judgement, tax officials may hesitate to override AI-generated decisions.

To mitigate these risks, AI oversight mechanisms must integrate a ‘human-in-the-loop approach,’ ensuring human officers remain actively involved in reviewing automated assessments (Mosqueira-Rey et al. 2023). A well-designed AI auditing system should simulate evaluative processes of human tax professionals while considering potential alternative determinations. By incorporating legal principles, administrative policies, and recognised bias sources, such oversight would enhance fairness and accuracy. As AI models evolve, these tools must be periodically updated to reflect legislative and judicial changes, preventing outdated legal interpretations from perpetuation.

The recommendation to use AI to supervise AI does not aim to replace human oversight, as required under the EU AI Act (Arts. 14 and 29), but to enhance it through a layered model that combines human and machine monitoring for greater scalability, consistency, and error detection. In high-volume domains like tax fraud detection, a secondary AI system can audit the primary system’s outputs, detect biases or anomalies, and verify compliance with legal thresholds—tasks that are often too complex or voluminous for continuous human review (Zhou et al. 2020; Lepri et al. 2018). This approach parallels AI assurance models, where layered systems are used to maintain algorithmic accountability. The secondary AI serves as a technical compliance tool, not a decision-maker, preserving human authority over final outcomes. Although there is a risk of error from the second AI, its role is to flag irregularities that may otherwise go unnoticed. This proposal aligns with emerging scholarship advocating for multi-layered safeguards in automated decision-making (Brundage et al. 2020; Selbst and Barocas 2018).

At the policy level, safeguards should also address imbalances in transparency. In contexts where AI tools are deployed without sufficient protection of fundamental rights, tax secrecy laws may need reconsideration (Kuzniacki et al. 2022). The differing interests of taxpayers and authorities necessitate nuanced approaches to AI-driven explanations. For taxpayers, systems must provide clear, accessible justifications, while for authorities, they should facilitate compliance and enforcement.

The obligation for governments to provide intelligible explanations of AI-informed decisions aligns with fundamental transparency principles. Where necessary, technical explanations should be available for independent experts to verify AI system coding and algorithmic logic, ensuring tax determinations and their rationales undergo rigorous scrutiny.

Ultimately, effective AI integration into tax administration requires robust, transparency-focused mechanisms safeguarding taxpayer rights while maintaining system efficiency. Independent AI-driven review systems would address critical risks associated with algorithmic opacity and automation bias, reinforcing procedural fairness and accountability. Given the increasing reliance on automated systems and the general lack of explainability in decisions, the absence of oversight poses significant concerns. Addressing these challenges requires proactive regulatory intervention at national and potentially international levels to ensure AI-driven tax systems operate within transparent, just legal frameworks.

## Conclusion

As AI continues to play an increasingly important role in tax administration, ensuring transparency, accountability, and fairness must be a priority. While AI-driven decision-making offers efficiency gains, it raises concerns about procedural fairness, due process, and algorithmic opacity. Governments should collaborate with legal scholars, technology specialists, professional bodies, and civil society stakeholders to design AI governance frameworks that uphold core taxpayer rights. Key concerns such as algorithmic transparency, data privacy, fairness, and accountability necessitate robust regulatory frameworks ensuring AI remains a tool for justice rather than unchecked automation.

The risks highlighted in cases such as Robodebt and Pintarich underscore the potential dangers of AI-driven decision-making when legal safeguards and human oversight are insufficient. These examples illustrate how overreliance on automation without appropriate procedural fairness mechanisms can erode public trust and lead to unjust outcomes. Addressing these risks requires commitment to transparency, ensuring taxpayers have access to

explanations for AI-generated determinations and meaningful avenues for review and appeal.

A key step in achieving this balance is establishing dedicated AI systems designed to explain and scrutinise tax authority determinations. This separate system would serve as an independent mechanism to interpret complex tax decisions, provide clear justifications, and allow taxpayers to contest automated assessments with meaningful oversight. By ensuring AI-generated tax determinations undergo both human and AI-led review, this approach enhances trust while safeguarding taxpayer rights.

The findings of this study highlight that while AI can improve efficiency and accuracy in tax administration, its adoption must be carefully aligned with legal and ethical principles to avoid undermining core taxpayer protections. From a theoretical standpoint, this work advances the understanding of the intersection between administrative law and AI governance. Scientifically, it contributes to the emerging body of interdisciplinary scholarship on algorithmic accountability and fairness. Practically, the analysis provides actionable insights for policymakers, regulators, and tax authorities seeking to design frameworks that balance innovation with integrity.

Nevertheless, this study has limitations. It primarily draws upon limited existing examples of case law and experiences, which may not fully capture the diversity of challenges faced in various jurisdictions with different legal traditions, tax systems, and levels of digital maturity. Future research should explore comparative perspectives across jurisdictions, assess the long-term impact of AI adoption on taxpayer trust, and investigate how emerging technologies such as explainable AI can be operationalised in real-world tax contexts.

Ultimately, the challenge lies in striking a balance between innovation and integrity. AI should serve as an enabler of tax justice, reinforcing legal protections rather than undermining them. By embedding theoretical, scientific, and practical insights into governance design, a future-proof tax administration system can embrace AI's efficiencies while embedding safeguards that uphold the rule of law, procedural fairness, and taxpayer rights. A governance framework integrating independent AI scrutiny, legal accountability, and transparent decision-making processes will be critical in maintaining public confidence and ensuring AI-driven tax determinations adhere to fundamental principles of fairness and justice.

### Data availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Received: 26 May 2025; Accepted: 13 October 2025;

Published online: 24 November 2025

### References

- Afield WE (2019) Social justice and the low-income taxpayer. *Vill L Rev* 64:347
- Alston P, Khawaja B, Riddell R (2019) Much ado about poverty: the role of a UN special rapporteur. *J Poverty Soc Justice* 27:423–429
- Amparo M, Ruiz G (2022) Fiscal transformations due to AI and robotization: where do recent changes in tax administrations, procedures and legal systems lead us? *Northwest J Technol Intellect Prop* 19:325
- Aslett J, Bruckner C, Coil C et al. (2024) Understanding artificial intelligence in tax and customs administration. IMF Tech Notes Man 2024/06. Washington, DC: International Monetary Fund
- ATO Our Charter (n.d.) <https://www.ato.gov.au/About-ATO/Commitments-and-reporting/ATO-Charter/Our-Charter> Accessed 22 October 2025
- Australian Human Rights Commission (2021) Human rights and technology. Final Report. Sydney NSW
- Australian National Audit Office (2024) Governance of artificial intelligence at the Australian Taxation Office. Auditor-General Report No. 26 of 2024–25. Canberra ACT
- Australian Taxation Office (2024a) How we use data and analytics. <https://www.ato.gov.au/about-ato/commitments-and-reporting/information-and-privacy/how-we-use-data-and-analytics>. Accessed 14 May 2025
- Australian Taxation Office (2024b) Inquiry into use of artificial intelligence systems by public sector entities. ATO Submission. Canberra ACT
- Bal A (2019) Ruled by algorithms: the use of black box models in tax law. *Tax Notes Int* 95:1347–1351
- Bardutzky S (2022) Duty to provide reasons. *Oxford Public International Law*. <https://opil.ouplaw.com/display/10.1093/law-oeul/law-oeul-e57>. Accessed 14 May 2025
- Barocas S, Hardt M, Narayanan A (2019) *Fairness and machine learning: limitations and opportunities*. MIT Press, Cambridge
- Bentley D (2022) Tax officer 2030: the exercise of discretion and artificial intelligence. *eJournal Tax Res* 20:72–100
- Binns R, Veale M (2021) Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR. *Int Data Priv Law* 11:319–344
- Bruckner C, Coil C (2024) AI and the modern tax agency: adopting and deploying AI to improve tax administration. American University Kogod Tax Policy Center, IBM Center for the Business of Government. Washington, D.C
- Brundage M et al. (2020) Toward trustworthy AI development: mechanisms for supporting verifiable claims. arXiv:2004.07213 <https://arxiv.org/abs/2004.07213>
- Brynjolfsson E, McAfee A (2017) *Machine, platform, crowd: harnessing our digital future*. W.W. Norton & Company, New York
- Coglianes C, Lehr D (2017) Regulating by robot: administrative decision making in the machine-learning era. *Georgetown Law J* 105:1147
- Cumberland T (2024) A look at the risks and opportunities for AI in tax administrations. OECD. AI Policy Observatory. Paris
- Denver University Law Review (2011) No taxation without representation: a British perspective on constitutional arrangements. *Denver Univ. Law Rev.* 88, 325
- Edwards L, Veale M (2017) Slave to the algorithm? Why a right to an explanation is probably not the remedy you are looking for. *Duke Law Technol Rev* 16(1):18–84
- EU AI Act. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024
- European Commission (2020) Fiscalis programme: supporting cooperation in the field of taxation. [https://ec.europa.eu/taxation\\_customs/fiscalis-programme\\_en](https://ec.europa.eu/taxation_customs/fiscalis-programme_en). Accessed 14 May 2025
- European Parliament (2025) EU AI Act: first regulation on artificial intelligence. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>. Accessed 20 July 2025
- Fink M, Finck M (2022) Reasoned A(I)ministration: explanation requirements in EU law and the automation of public administration. *Eur Law Rev* 47:376–392
- Gorski L et al (2024) Exploring explainable AI in the tax domain. *Artif Intell Law*. <https://doi.org/10.1007/s10506-024-09395-w>
- IBDF (2020) Observatory on the protection of taxpayers' rights. The IBFD yearbook on taxpayers' rights 2020. <https://www.ibfd.org/ibfd-academic/publications/observatory-protection-taxpayers-rights-optr>. Accessed 21 July 2025
- Internal Revenue Service (2022) IRS strategic plan: fiscal years 2022–2026. <https://www.irs.gov/pub/irs-pdf/p3744.pdf>. Accessed 14 May 2025
- IRS taxpayer bill of rights <https://www.irs.gov/taxpayer-bill-of-rights#:~:text=The%20Right%20to%20Be%20Informed,Fair%20and%20Just%20Tax%20System>
- Kollruss T (2023) A hybrid entity structure for tax optimization of foreign direct investment in the U.S. *J Corp Acc Finance* 35(2):153
- Kollruss T (2025b) Is the tax calculation method for exemptions with progression contrary to EU law? Quantitative and formal-analytical analysis. *Financ Innov* 11:89. <https://doi.org/10.1186/s40854-024-00706-3>
- Kollruss T (2025a) The (social) tasks of business tax research and the binding effect of the statutory tax burden decision. *Account Econ Law A Conviv* 15(2):325–346. <https://doi.org/10.1515/acl-2022-0089>
- Kollruss T (2025c) Group taxation and EU law. *Liverp Law Rev* 46:181–193. <https://doi.org/10.1007/s10991-025-09382-4>
- Kuzniacki B et al (2022) Towards explainable artificial intelligence (XAI) in tax law: the need for a minimum legal standard. *World Tax J* 14:573–616
- Lepri B, Oliver N, Letouze E, Pentland A, Vinck P (2018) Fair, transparent, and accountable algorithmic decision-making processes. *Philos Technol* 31(4):611–627
- McCarty LT (1977) Reflections on TAXMAN: an experiment in artificial intelligence and legal reasoning. *Harv Law Rev* 90:837–893
- Mohun J, Roberts A (2020) Cracking the code: rulemaking for humans and machines. OECD Working Papers on Public Governance No. 42
- Mosqueira-Rey E et al. (2023) Human-in-the-loop machine learning: a state of the art. *Artif Intell Rev* 56:3005–3054

- Myburg J, Schutte D (2025) SMMEs' business mindset on tax compliance in North West: perceptions and challenges. *Southern Afr J Entrep Small Bus Manag* 17(1):1-13
- Nathwani K (2024) Artificial intelligence in automated decision-making in tax administration: the case for legal, justiciable and enforceable safeguards. Institute for Fiscal Studies. London
- Nederlands Juristen Comite Voor de Mensenrechten [NJCM] et al v State of the Netherlands, The Hague District Court Case No. C/09/550982 / HA ZA 18-388 (2020)
- Ng YF, Goh YJ, Tan N et al. (2020) Revitalising public law in a technological era: rights, transparency and administrative justice. *Univ N South Wales Law J* 43:1041
- OECD (2022) Framework for the classification of AI systems. OECD Publishing, Paris
- OECD (2025) Tax administration digitalisation and digital transformation initiatives. OECD Publishing, Paris
- Pintarich v Deputy Commissioner of Taxation (2018) FCAFC 79; 262 FCR 41; 108 ATR 31
- Royal Commission into the Robodebt Scheme (2023) Report. Commonwealth of Australia. Canberra
- Saba CS, Monkam NF (2024) Leveraging the potential of artificial intelligence (AI) in exploring the interplay among tax revenue, institutional quality, and economic growth in the G-7 countries. *AI Soc.* <https://doi.org/10.1007/s00146-024-01885-4>
- Selbst AD, Barocas S (2018) The intuitive appeal of explainable machines. *Fordham Law Rev* 87(3):1085-1139
- United Nations (2019) Digital technology, social protection and human rights: report. <https://www.ohchr.org/en/calls-for-input/digital-technology-social-protection-and-human-rights-report>. Accessed 14 May 2025
- Veale M, Borgesius FZ (2021) Demystifying the draft EU Artificial Intelligence Act. *Comput Law Rev Int* 22(4):97-112
- Wachter S, Mittelstadt B, Floridi L (2017) Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *Int Data Priv Law* 7:76-99
- Werfel D (2023) Written testimony of Daniel Werfel, IRS Commissioner before the House Oversight and Accountability Committee. <https://www.irs.gov/newsroom/written-testimony-of-daniel-werfel-commissioner-internal-revenue-service-before-the-hoac-subcommittees-on-government-operations-and-the-federal-workforce-and-health-care-and-financial-services-and-irs-operations>. Accessed 14 May 2025
- Witten IH, Frank E, Hall MA (2016) Data mining: practical machine learning tools and techniques, 4th edn. Morgan Kaufmann, Burlington
- Zalnieriute M, Moses LB, Williams G (2021) The rule of law and automation of government decision-making. *Mod Law Rev* 84(3):425-457
- Zheng Q, Xu Y, Liu H, Shi B, Wang J, Dong B (2024) A survey of tax risk detection using data mining techniques. *Engineering* 34:43-59
- Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, Wang L, Li C, Sun M (2020) Graph neural networks: a review of methods and applications. *AI Open* 1:57-81

## Author contributions

The corresponding author contributed to and supervised this work.

## Competing interests

The author declares no competing interests.

## Ethical approval

This study did not involve human participants or identifiable personal data. It is based solely on publicly available legal sources and academic literature. In line with the Australian National Statement on Ethical Conduct in Human Research (NHMRC, 2018), such research is exempt from ethics review.

## Informed consent

This research did not involve human participants or the collection of personal data; therefore, informed consent was not required.

## Additional information

**Correspondence** and requests for materials should be addressed to Evgeny Guglyuyatyy.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025