ALGORITHMIC SYSTEMS AND DEMOCRATIC OVERSIGHT IN PUBLIC AUDITING

KAMU DENETİMİNDE ALGORİTMİK SİSTEMLER VE DEMOKRATİK DENETİM

Hamza ATEŞ[1]

ABSTRACT

This article examines the transformative impact of artificial intelligence (AI) on public auditing, addressing its normative, procedural, and institutional dimensions. As AI systems are increasingly deployed for functions such as fraud detection, performance monitoring, and predictive compliance, they are reshaping the Weberian foundations of public auditing—traceability, procedural accountability, and human judgment. Drawing on public administration theory, science and technology studies, and the literature on algorithmic governance, the study investigates how algorithmic opacity, epistemic asymmetries, and machine-based risk logics undermine the democratic legitimacy of the audit function. Through case studies from the Netherlands, Estonia, Brazil, and the United States, it highlights both the opportunities and risks of algorithmic auditing in practice. The article proposes a normative framework built upon four guiding principles—transparency, auditability, contestability, and institutional ethics—to ensure accountable and just integration of AI into public oversight. Rather than advocating for the wholesale rejection or uncritical embrace of AI, the article argues for a deliberate rethinking of human–machine relations in public administration and auditing, contending that only in this way can the integrity of democratic institutions be safeguarded in the algorithmic age.

1- Prof. Dr., İstanbul Medeniyet Üniversitesi, ates.hamza@gmail.com, ORCID: 0000-0003-0975-0062.

ÖZ

Bu makale, yapay zekânın (YZ) kamu denetimi (audit) üzerindeki dönüştürücü etkisini, normatif, prosedürel ve kurumsal boyutlarıyla ele almaktadır. Yolsuzluk tespiti, performans izleme ve öngörücü uyumluluk gibi işlevlerde giderek daha fazla kullanılan YZ sistemleri, kamu denetiminin Weberyen temelleri olan izlenebilirlik, usule uygun hesap verebilirlik ve insan yargısı ilkelerini yeniden şekillendirmektedir. Kamu yönetimi kuramı, bilim ve teknoloji çalışmaları ile algoritmik yönetişim literatüründen yararlanan bu analiz, algoritmik kapalılığın, bilgi eşitsizliklerinin ve makine-temelli risk mantıklarının, denetim işlevinin demokratik meşruiyetini hangi yollarla zayıflattığını araştıtmaktadır. Hollanda, Estonya, Brezilya ve ABD'den vaka örnekleri üzerinden yapılan değerlendirmeler, algoritmik denetimin pratikteki imkânlarını ve risklerini ortaya koymaktadır. Makale, kamu gözetiminde YZ'nin hesap verebilir ve adil biçimde bütünleşmesini sağlamak üzere dört ilkeye dayanan normatif bir çerçeve önermektedir: şeffaflık, denetlenebilirlik, itiraz edilebilirlik ve kurumsal etik. Makalede, YZ'yi toptan reddetmek ya da sınırsızca benimsemek yerine, kamu yönetimi ve denetiminde insan–makine ilişkilerinin yeniden düşünülmesi gerektiği savunulmakta ve demokratik kurumların bütünlüğünü korumanın ancak bu şekilde mümkün olabileceği anlatılmaktadır.

Keywords: Algorithmic auditing, Public accountability, Artificial intelligence in auditing, Administrative ethics.

Anahtar Kelimeler: Algoritmik denetim, Kamu hesap verebilirliği, Denetimde yapay zekâ, İdari etik.

INTRODUCTION

The integration of artificial intelligence (AI) into public administration represents not only a technological innovation but also a fundamental transformation of the principles and practices that underpin state functions such as public auditing. Traditionally, public auditing has been associated with neutrality, legality, and procedural accountability, serving as a key mechanism for safeguarding the integrity of democratic governance (Power, 1997; Weber, 1978; Afzal, 2025: 2; Pérez-Durán, 2024: 1163–1165). Auditors were expected to ensure that public resources were used lawfully and effectively, with clear lines of responsibility and transparent procedures. Scholars have emphasized that algorithmic governance is not only a technical process but also a political and institutional reconfiguration of authority, requiring systematic research into its democratic implications (Danaher et al., 2017: 3–5).

The growing adoption of AI systems in audit processes challenges these foundations. Machine learning and data-driven analytics are increasingly employed to detect fraud, monitor performance, and predict risks with unprecedented speed and scale. This trend is often framed as administrative modernization, promising efficiency and real-time responsiveness. Yet the very techniques that underpin AI—probabilistic modeling, adaptive algorithms, and predictive analytics—introduce new complexities. Unlike traditional audits, which operated retrospectively and could trace decisions to established rules, algorithmic systems often function as opaque "black boxes", difficult to interpret even by their creators (Burrell, 2016; Pasquale, 2015). This opacity raises questions of explainability, contestability, and fairness, thereby unsettling the democratic legitimacy of audits (Ananny & Crawford, 2018).

AI's predictive capabilities also echo earlier debates on continuous auditing (CA). Emerging in the early 1990s, CA sought to use technology for high-frequency and real-time monitoring (Vasarhelyi & Halper, 1991; Alles et al., 2006). While AI realizes many of CA's ambitions—ongoing surveillance, anomaly detection, and anticipatory oversight—it differs in its reliance on probabilistic, self-learning models. This shift risks collapsing the distinction between monitoring and enforcement, amplifying concerns about discretion, bias, and due process (Murikah, 2024: 6–8; Mahroof et al., 2025: 2–4). The convergence of AI and CA thus exemplifies how technological innovation can both extend and destabilize established oversight logics. These developments highlight a central tension: whether trust in public auditing should rest on technical accuracy or on procedural and democratic accountability. As AI reshapes the epistemic and institutional basis of audits, it compels a rethinking of how authority and responsibility are defined in governance, increasingly requiring participatory mechanisms that bring publics into the audit process (Eslami et al., 2025: 5–7).

In approaching this inquiry, the article draws on interdisciplinary scholarship spanning public administration, science and technology studies, legal theory, and political philosophy. It treats AI not merely as a tool but as a socio-technical assemblage that reshapes institutional practices, epistemic assumptions, and political relationships. Accordingly, it calls for a conceptual reframing of AI auditing as a site of contested governance—a domain where technical innovation, democratic norms, and public ethics intersect,

sometimes productively and sometimes in tension. Importantly, the article rejects both technological determinism and uncritical enthusiasm. It does not assume that AI will necessarily enhance or erode accountability but asks under what conditions, through what institutional designs, and with what normative commitments AI can contribute to the public good. By interrogating the changing logics of public auditing in the age of AI, it seeks to illuminate the possibilities and perils of administrative machines in democratic governance. Finally, the article contributes to a growing body of scholarship calling for algorithmic accountability as a new frontier in public ethics. As AI systems become embedded in the core functions of the state, including taxation, welfare distribution, policing, and procurement, the audit function becomes a crucial site for ensuring that technological power is exercised with legitimacy, restraint, and justice. In this endeavor, public auditing must evolve—not by abandoning its principles but by rearticulating them in ways that are adequate to the demands of the digital state.

1. THEORETICAL FOUNDATIONS: FROM BUREAUCRATIC RATIONALITY TO ALGORITHMIC GOVERNANCE

The transformation of public auditing through AI cannot be understood without first situating it within the broader conceptual evolution of state rationality. The emergence of algorithmic systems in governance marks a significant shift in how administrative authority is conceptualized, exercised, and justified. To grasp this shift, it is essential to revisit the foundational principles of bureaucratic legitimacy as articulated by Max Weber, and consider contemporary calls for procedural fairness and public engagement in algorithmic decision-making (Decker et al., 2025: 4–7; Parviainen et al., 2025: 3–5) and to contrast them with the emergent logic of algorithmic governance, which reconfigures key dimensions of decision-making, accountability, and transparency. This aligns with calls to treat algorithmic governance as a distinct research agenda that integrates legal, ethical, and institutional dimensions (Danaher et al., 2017: 7–10).

Max Weber's ideal-type of bureaucracy was premised on the rational-legal authority of the modern state, where administrative legitimacy derived from a system of rules, hierarchical organization, and impersonal decision-

making (Weber, 1978: 956–958). Auditing, in this context, is an extension of rational-legal authority: it evaluates whether actions taken by officials conform to rules, laws, and budgetary frameworks. The audit is both a technical and a moral operation, securing the public's trust through structured oversight mechanisms grounded in procedural rationality (Power, 1997: 21–23). For Weber, auditability is embedded in the very architecture of bureaucracy. Administrative actions must be documented, decisions must be reasoned, and responsibilities must be traceable (Weber, 1978: 988). This traceability underpins the audit function: it allows for ex-post verification and creates the conditions for accountability.

However, as governance structures became more complex, the Weberian framework faced challenges. Decentralization, privatization, and globalization created conditions in which traditional ex-post audits appeared increasingly inadequate. This "audit explosion" (Power, 1997: 1–4) generated demand for new instruments of oversight. Here, the transition to algorithmic governance is not merely a technological shift but a transformation in administrative rationality. AI systems, particularly machine learning models, embody a move from rule-based reasoning to pattern-based prediction, thus altering legitimacy from procedural adherence to performance-based validation (Pasquale, 2015: 8). This movement resonates with the conceptual lineage of continuous auditing (Vasarhelyi & Halper, 1991; Alles et al., 2006), where oversight is envisioned as ongoing and anticipatory rather than episodic. Situating AI in relation to both Weberian bureaucracy and continuous auditing clarifies the hybrid nature of the emerging paradigm.

In this context, the rise of algorithmic governance represents not merely a technological innovation but a reconstitution of administrative rationality itself. AI systems do not follow static rules but learn patterns from large datasets, optimizing decisions through probabilistic modeling rather than fixed legal frameworks. This shift entails a movement from rule-based reasoning to pattern-based prediction, from process legitimacy to output optimization. Algorithmic systems promise enhanced efficiency, scalability, and objectivity, aligning with longstanding bureaucratic ideals. Yet the logic they introduce differs fundamentally from Weberian proceduralism. Instead of deriving legitimacy from rule adherence, algorithmic decisions are often justified through their performance—accuracy rates, predictive validity, or

error reduction (Pasquale, 2015: 8). This shift from legal-rational justification to techno-functional validation marks a profound transformation in how authority is exercised and evaluated.

Crucially, algorithmic systems often operate in opaque or non-intelligible ways. Machine learning models, particularly deep learning architectures, may generate decisions whose internal processes cannot be easily reconstructed or explained—a phenomenon sometimes referred to as the black box problem (Burrell, 2016: 1–3). This challenges the auditability of decisions, undermining the very premise of transparency that legitimized earlier bureaucratic models. The logic of explainability is replaced by a logic of performance metrics, which may be inaccessible or unintelligible to external auditors and lay citizens alike. Moreover, algorithmic governance alters the temporality and locus of decision-making. Traditional bureaucratic decisions are discrete, traceable events, embedded within hierarchical chains of command. Algorithmic decisions, by contrast, are often continuous, automated, and distributed—embedded in systems that constantly update and recalibrate themselves based on new data inputs (Yeung, 2018: 508–509). This undermines the possibility of ex-post evaluation in any meaningful sense, as the decision space is no longer static but dynamic.

This new logic also reconfigures the nature of discretion in public administration. Whereas bureaucratic discretion was historically framed as a human faculty requiring judgment and moral reasoning within the boundaries of legal norms, algorithmic discretion is emergent and embedded. It is not located in any single official but encoded in the design, training, and deployment of systems. As a result, traditional models of responsibility attribution—who decided what and why—are rendered inadequate (Wieringa, 2020: 395–397). The conditions of "auditability," central to Weberian legitimacy, are displaced by conditions of traceability without intelligibility. This epistemic transformation extends beyond technical systems into institutional cultures. Public agencies increasingly adopt data-driven management styles, where decision quality is inferred from quantifiable outcomes rather than procedural compliance. Risk assessment tools, predictive analytics, and automated flagging mechanisms shift the role of the auditor from a reviewer of past actions to a monitor of behavioral trends (Zuboff, 2019: 214–216). This positions auditing within a new paradigm of anticipatory governance—one that aims not to detect failure but to prevent it, often by intervening in advance of wrongdoing.

Such a shift raises profound normative concerns. Anticipatory auditing can easily turn into surveillance, eroding the boundary between compliance and control. It may also reinforce existing inequalities if the data used to train algorithms reflect historical biases or institutionalized discrimination (Eubanks, 2018: 72–75). The legitimacy of decisions thus becomes contingent not on formal rule conformity but on statistical correlations and operational metrics. In sum, the emergence of AI in public auditing signifies a movement from a model of governance grounded in legal rationality to one structured by algorithmic rationality. This transformation redefines key elements of auditing—such as accountability, transparency, discretion, and trust—posing challenges to both traditional public administration and contemporary democratic theory. As auditing becomes increasingly automated, the normative principles that once sustained its legitimacy must be critically re-examined and rearticulated for a new institutional terrain.

2. AI AND THE AUDIT FUNCTION: EFFICIENCY, RISK, AND SURVEILLANCE

The integration of AI into public auditing closely aligns with and extends the concept of continuous auditing (CA). Continuous auditing, first theorized in the early 1990s (Vasarhelyi & Halper, 1991), refers to audit processes enabled by information technologies that allow for high-frequency or even real-time evaluation of financial and operational data. In contrast to periodic audits, which occur at fixed intervals, continuous auditing provides near-instantaneous assurance and monitoring. Subsequent scholarship (Alles et al., 2006) emphasized its potential to enhance anomaly detection, accelerate risk identification, and improve managerial responsiveness. AI amplifies these capacities by automating anomaly detection, risk scoring, and predictive compliance on an unprecedented scale. Machine learning systems can process vast streams of financial and operational data in real time, flagging irregularities that warrant further human investigation. Predictive analytics also enable auditors to anticipate risks before they materialize, aligning closely with the preventive orientation of continuous auditing.

Despite these overlaps, AI introduces distinctive features into continuous auditing practices. First, while classical CA frameworks relied on deterministic rule sets and automated controls, AI systems leverage probabilistic modeling and adaptive learning. This shift introduces new opportunities for pattern recognition but also exacerbates concerns about algorithmic opacity and explainability. Second, AI-driven continuous auditing raises normative questions regarding surveillance and discretion. The ability to monitor continuously and predictively can erode the boundary between audit oversight and managerial control, raising due process concerns similar to those observed in predictive policing (Eubanks, 2018: 95–98).

Comparing AI-enabled auditing with established CA practices therefore highlights both continuity and rupture. On the one hand, AI realizes many of the promises envisioned by continuous auditing scholars: efficiency, timeliness, and risk anticipation. On the other hand, it transforms the epistemic and ethical terrain of auditing, shifting from rule-based monitoring toward predictive governance. As such, AI-based continuous auditing must be situated within broader debates about democratic legitimacy, accountability, and the balance between efficiency and rights. On the other hand, AI is rapidly reconfiguring the landscape of public auditing, introducing new techniques, rationalities, and institutional expectations. At the core of this transformation lies a fundamental redefinition of what constitutes effective and legitimate audit practice. Where traditional audits centered on retrospective analysis, rule compliance, and human judgment, AI-enabled audits emphasize speed, pattern recognition, and anticipatory control. This shift is often presented as a natural evolution—an enhancement of the audit's core function. Yet, beneath the surface lies a complex interplay of epistemic, ethical, and political dynamics.

AI's entry into public auditing is most visibly manifested in its applications for fraud detection. Using machine learning algorithms trained on large datasets of financial transactions, procurement records, and tax filings, audit institutions can now identify anomalies or suspicious behaviors that deviate from normative patterns (OECD, 2019: 11–14; Yavuz, 2024). These tools surpass human capacity in terms of speed and scale, enabling near real-time flagging of irregularities across massive datasets. For instance, neural networks can detect subtle correlations between vendor behavior, invoice structures, and payment patterns that might signal corruption or collusion—insights that

would elude conventional audit methods. This expansion of analytical capacity has led to a growing reliance on performance measurement as a core audit function. Performance audits historically assessed whether public agencies were using resources effectively, efficiently, and economically. AI systems can automate this task by continuously evaluating key performance indicators (KPIs) through data mining and real-time analytics. In many jurisdictions, public managers now interface with dashboards populated by AI-generated visualizations, predictive metrics, and alerts. While such tools promise enhanced managerial oversight, they also risk reducing complex policy goals to reductive, quantifiable benchmarks—what Strathern (1997: 308) famously cautioned against: "When a measure becomes a target, it ceases to be a good measure".

A further development is the deployment of AI in predictive compliance—a technique that seeks to identify and preempt non-compliant behavior before it occurs. These systems use probabilistic modeling to assess the risk that an entity (a public contractor, a welfare recipient, or even a municipal office) may engage in misconduct. In doing so, they generate risk scores that can trigger intensified scrutiny, audits, or interventions. This preventive logic signals a profound shift in the orientation of auditing—from ex-post accountability (evaluating actions after the fact) to ex-ante monitoring (anticipating risk before action). The implications of this shift are profound. Traditionally, accountability was rooted in the retrospective reconstruction of decision processes. It required documentation, procedural fairness, and the possibility of appeal. AI-enabled audits, by contrast, often act upon inferences and patterns rather than demonstrated violations. This raises concerns about due process, particularly when interventions—such as benefit suspension or contract denial—are based on risk predictions rather than proven infractions (Eubanks, 2018: 95–98). The audit function thus begins to resemble a form of surveillance, where the subject is monitored not for what they have done, but for what they might do.

This anticipatory turn in auditing mirrors broader transformations in governance, often described as part of a move toward risk-based regulation (Black, 2005: 3–6). In this model, oversight bodies allocate attention and resources based on probabilistic risk rather than equal or random audit selection. AI enhances this model by refining risk scores with increased

granularity and predictive power. While this appears efficient from a resource allocation standpoint, it also introduces the danger of discriminatory automation. Communities or individuals whose behaviors statistically correlate with past violations may be disproportionately targeted, even if such patterns are the result of structural disadvantage rather than individual malfeasance. These systems also present challenges to the principle of audit traceability. AI-generated decisions may rely on complex models that are not easily interpretable by auditors themselves. If the reasoning behind a flagged irregularity is opaque, auditors must either trust the model blindly or override it without sufficient justification—both scenarios undermining the epistemic integrity of the audit process (Burrell, 2016: 4–6). In practice, this can lead to a double-bind: either excessive reliance on algorithmic outputs (technocratic overreach) or refusal to engage with them altogether (technophobia).

Furthermore, the adoption of predictive auditing tools alters the power dynamics within public administration. Traditional audits maintained a distinction between those who governed and those who were scrutinized. AI-enabled systems blur this line, as public servants themselves become subjects of automated monitoring. Time-on-task, email activity, or keystroke patterns may be used to assess bureaucratic productivity or detect anomalies. The auditor becomes an omnipresent function, diffused across layers of surveillance infrastructure rather than instantiated in a specific institutional actor. This surveillance logic can undermine the relational trust that auditing systems aim to foster. When civil servants, vendors, or citizens feel that they are being preemptively judged by inscrutable algorithms, their willingness to cooperate, report irregularities, or improve systems may diminish. As O'Neil (2016: 142–144) has shown, opaque and punitive algorithmic systems often generate compliance without consent—obedience born not of legitimacy but of fear. Such dynamics are antithetical to the broader goals of responsive and participatory public administration.

Proponents of AI-based auditing argue that these tools reduce corruption, enhance detection, and introduce objectivity. However, this technocratic promise must be weighed against the normative costs of automation. The substitution of algorithmic suspicion for human deliberation risks shifting the audit function from one of public accountability to one of instrumental control. Auditing, in this model, no longer reveals wrongdoing so

much as it manufactures a continuous condition of potential guilt. This also challenges the institutional culture of audit institutions. Many public auditors view their role not merely as fault-finders, but as partners in improving governance—providing recommendations, building institutional capacity, and fostering trust. AI-based auditing, with its emphasis on detection and risk management, may marginalize these functions, reducing the auditor's role to that of a data technician or algorithmic supervisor.

Moreover, efficiency itself becomes a contested concept in this context. While AI systems undoubtedly increase the speed and scope of analysis, they may do so at the expense of interpretive nuance and contextual understanding. Auditors must grapple with whether efficient detection equals just auditing, and whether predictive analytics can ever substitute for procedural fairness and human judgment. The ideal of auditing as a "reason-giving" practice (Bovens, 2007: 451) is at risk when the reasons behind actions are machine-derived, probabilistic, and inaccessible. In sum, the integration of AI into public auditing is not merely a technical upgrade—it constitutes a reconfiguration of the audit function itself. The movement from retrospective verification to real-time prediction represents a shift in the very temporal, epistemic, and moral logics of auditing. While enhanced efficiency and risk management are important goals, they must not come at the expense of core democratic principles: transparency, accountability, and justice. Future sections of this article will explore how these tensions can be addressed through institutional reform, participatory oversight, and the development of algorithmic accountability frameworks.

3. TRUST, TRANSPARENCY, AND THE BLACK BOX PROBLEM

Trust has always been foundational to the legitimacy of public auditing. Citizens, legislators, and institutions alike must trust that auditors operate impartially, transparently, and competently in evaluating the use of public resources. In the Weberian tradition, such trust was cultivated through procedural safeguards, written documentation, and the clear attribution of responsibility to identifiable officials. Trust in this sense was not a mere sentiment but a structured expectation that public auditors could reconstruct administrative decisions and justify them with reference to established rules

and standards (Bovens, 2007: 451–452). The expansion of AI into audit practices, however, unsettles these foundations. As decisions increasingly stem from opaque and probabilistic systems rather than from human deliberation, the conditions under which trust is generated and sustained undergo a profound transformation.

This shift raises important questions about the basis of public trust. Is legitimacy anchored in the technical accuracy of AI models, or in the moral and procedural accountability of human auditors? Studies suggest that citizens often demand not only effective outcomes but also fairness and transparency in the processes that govern them (Bovens, 2007: 452–454; Yener at al., 2025: 13). As such, the introduction of AI threatens to replace an ethos of reason-giving with one of statistical efficiency, creating a gap between technical performance and democratic legitimacy. This tension suggests that trust in public auditing cannot be reduced to confidence in machine outputs but must also include the assurance that decisions are understandable, challengeable, and consistent with societal values. At the heart of this tension lies the black box problem—the difficulty, or in many cases the impossibility, of understanding how complex algorithmic systems produce their outputs (Burrell, 2016: 3–5). Unlike rule-based auditing, where each decision can be traced back to a regulatory or legal principle, machine learning systems often function as opaque infrastructures that process massive datasets through layers of statistical correlation. This opacity is not merely a technical inconvenience but a structural characteristic of contemporary AI, where even developers may struggle to fully explain why a model produces a given output (Pasquale, 2015: 19).

The challenges of explainability in AI auditing resonate strongly with earlier debates on continuous auditing, where the move toward real-time monitoring also raised concerns about over-reliance on automated processes without adequate human interpretation (Vasarhelyi & Halper, 1991). In both cases, efficiency gains risk undermining transparency and accountability if the processes underlying decisions are inaccessible to scrutiny. As Ananny and Crawford (2018: 975–976) argue, explainability in algorithmic governance is not solely a technical matter but a democratic requirement: citizens and officials must be able to grasp, at least in principle, how evaluative judgments are being made. Emerging approaches in explainable AI (XAI) aim to address

these concerns, offering methods for visualizing model logic or generating post hoc rationales (Doshi-Velez & Kim, 2017: 4–7). Yet, these explanations often remain intelligible only to specialists, doing little to address the broader public's demand for understandable justifications. In the context of public auditing, where accountability to citizens is paramount, the limits of XAI raise urgent questions about whether the black box can ever be rendered sufficiently transparent to meet democratic standards.

The absence of clear explanations creates epistemic asymmetries—unequal distributions of knowledge and interpretive capacity between those who design and control algorithms, those who operate within bureaucratic systems, and those who are subjected to algorithmic decisions (Pasquale, 2015: 8–10; Wieringa, 2020: 398–400; Damar et al., 2024: 13). These asymmetries are not merely technical; they have profound political consequences. When auditors lack the expertise to interrogate algorithmic outputs, they risk deferring uncritically to the authority of machine predictions, thereby hollowing out their own evaluative role. For affected citizens, epistemic asymmetries translate into a lack of meaningful recourse: individuals flagged by an algorithm may not understand the basis of the decision, nor possess the resources to contest it. Such asymmetries erode the democratic legitimacy of public auditing. The traditional model presumes a clear chain of responsibility—officials can be held accountable for their actions, and citizens can demand explanations grounded in law. Algorithmic discretion, however, disperses agency across developers, data scientists, vendors, and bureaucrats. As scholars such as Yeung (2018: 507–509) and Mökander, Axente, and Floridi (2021: 62–64) note, this distribution makes it difficult to identify who should be answerable when harms occur. In effect, accountability becomes diluted, undermining the very function of auditing as a guardian of democratic integrity.

To address these challenges, technical and institutional initiatives have emerged. Explainable AI (XAI) seeks to make algorithmic systems more interpretable by providing accessible accounts of how outputs are generated (Doshi-Velez & Kim, 2017). Algorithmic audit trails—records of data sources, model versions, and decision pathways—aim to restore traceability and verifiability to automated processes. Recent OECD (2022) and World Bank (2020) reports emphasize the importance of embedding such technical safeguards into public financial management systems to sustain trust. Yet,

transparency alone is insufficient. Participatory algorithmic governance offers a complementary pathway. Citizen panels, multi-stakeholder oversight committees, and deliberative forums can democratize algorithmic oversight, ensuring that explanations are not only technically valid but also socially intelligible and normatively meaningful (Ziewitz, 2019: 4–6). Moreover, the growing literature on AI ethics emphasizes that governance must go beyond compliance to include equity, fairness, and contextual sensitivity (Mökander et al., 2021: 60–62; Polat, 2024: 397). Without such mechanisms, AI-driven continuous auditing risks degenerating into a regime of technocratic control, where efficiency is privileged at the expense of justice and participation.

In short, trust in AI-enabled auditing systems depends not only on technical accuracy but also on institutional capacities for explanation, contestation, and ethical oversight. The black box problem and the resulting epistemic inequalities are not merely technical hurdles but fundamental challenges to democratic legitimacy.

4. CONTEXTUAL REALITIES: TRANSLATING ALGORITHMIC LOGIC INTO PRACTICE

While theoretical and normative analysis is indispensable in understanding the transformation of public auditing under AI, it must be complemented by empirical reflection. Examining concrete institutional cases reveals the tensions between promise and practice, innovation and opacity, and efficiency and equity. This section presents selected cases from jurisdictions where AI has been introduced into audit or audit-adjacent functions. These cases highlight both the potential of algorithmic auditing and the risks of its premature, opaque, or unaccountable deployment.

Before turning to specific case studies, it is important to clarify the methodological approach underpinning this analysis. Four cases were selected -the Netherlands, Estonia, Brazil, and the United States- because they represent diverse governance contexts and distinct trajectories of AI adoption in public auditing. The Netherlands and Estonia were chosen as European pioneers: the former due to its prominent childcare benefits scandal involving algorithmic risk models, and the latter because of its globally recognized e-governance infrastructure. Brazil offers insights from a Global South

democracy experimenting with municipal-level AI auditing initiatives, while the United States demonstrates the role of a federal-level audit body (GAO) in developing standards for AI accountability. The analysis draws primarily on official audit reports, government publications, legislative documents, and secondary scholarly literature published between 2018 and 2024. These sources were systematically reviewed to identify how AI was implemented, what governance mechanisms were applied, and how democratic legitimacy was challenged or reinforced. Each case contributes to the article's central argument by illustrating both the opportunities and risks of algorithmic auditing in real-world settings, thereby grounding the theoretical analysis in empirical evidence.

In one of the most high-profile responses to algorithmic risk in the public sector, the Dutch Court of Audit (Algemene Rekenkamer) has taken significant steps to examine the use of algorithms in central government agencies. Following the Dutch childcare benefits scandal (Toeslagenaffaire) -in which thousands of families, many with migrant backgrounds, were wrongly accused of fraud by an automated risk model- the Court began to systematically audit the government's use of algorithms for fairness, transparency, and legality (the Netherlands Court of Audit, 2021: 6–9). The model used by the Dutch Tax Administration was found to be non-transparent, discriminatory, and poorly governed, relying on sensitive variables such as dual nationality to flag risks without sufficient human oversight.

This case underscored the importance of algorithm governance as part of financial and administrative oversight. The Court emphasized that algorithms used in high-stakes public decisions must be auditable, explainable, and contestable. Importantly, it called for better documentation of model logic and more robust risk classification governance frameworks. However, the scandal also revealed the institutional lag between AI deployment and oversight mechanisms. By the time the issue reached public attention, significant harm had already occurred, and legal redress remained difficult. The Netherlands' case demonstrates that even in technically advanced and institutionally robust democracies, the use of AI in auditing functions can entrench bias and undermine trust if not governed proactively. It also illustrates the emerging role of courts of accounts not only as financial auditors but as guardians of algorithmic accountability.

Estonia, widely considered a pioneer of digital governance, has experimented with AI across various public domains, including auditing and service delivery. The Estonian National Audit Office has endorsed the use of predictive analytics to detect procurement anomalies, monitor contract performance, and anticipate irregularities in public spending. These systems are embedded within broader e-governance infrastructures such as the X-Road data exchange layer, which connects state registries and allows real-time cross-referencing of procurement, tax, and company data. While Estonia's model is often cited as a success case of digital trust, it is not without its tensions. The emphasis on automation and cross-sectoral data sharing has increased surveillance capacities and blurred boundaries between internal auditing and administrative policing. Moreover, the opacity of the algorithms used, particularly their risk-scoring mechanisms, remains a concern—both for auditors and for civil society groups advocating for algorithmic transparency. Notably, Estonian audit professionals have voiced concern that human capacity to interpret algorithmic results is declining, especially as model complexity increases. This raises questions about epistemic dependency: can auditors meaningfully contest or calibrate what they do not fully understand? Estonia thus illustrates both the strengths and the fragilities of AI-based auditing in high-trust, high-automation governance settings.

In São Paulo, the Municipal Comptroller General's Office (CGM-SP) launched a pilot program using machine learning to detect irregularities in municipal contracts and employee reimbursements. The system, developed in partnership with university researchers, analyzes historical financial and procurement data to flag outliers and rank audit priorities. While the program has shown early success in identifying previously undetected anomalies—such as ghost employees or price collusion—it has also sparked criticism about due process and over-reliance on correlation-based flags. Whistleblower protection groups raised concerns that algorithmic risk scores are being treated as evidence of wrongdoing, rather than as investigative leads. In some instances, flagged individuals were suspended before any human-led investigation had occurred. This case exemplifies the tension between efficiency and procedural fairness. By shifting the locus of suspicion upstream—from confirmed violations to probabilistic inferences—the system risks penalizing individuals without legal basis. São Paulo's experience suggests the need to carefully distinguish between risk prediction as a managerial tool and judgment as a legal and ethical responsibility.

The U.S. Government Accountability Office (GAO) has taken a leading role in setting methodological standards for auditing AI systems used in federal agencies. Recognizing the risks posed by opaque and biased algorithms, I released its "AI Accountability Framework", which provides a structured approach for evaluating transparency, reliability, privacy, and fairness in government AI applications (GAO, 2021: 2–4). While the GAO does not yet use AI tools extensively in its own audits, it has been proactive in articulating the conditions under which AI deployment is audit-worthy. GAO's framework encourages agencies to maintain clear documentation, conduct impact assessments, and build cross-functional governance teams that include ethicists, legal experts, and technical staff. Importantly, it advocates for public involvement and stakeholder engagement—marking a normative commitment to democratic oversight rather than mere technical compliance. Though still in an early stage of operationalization, GAO's initiative exemplifies how auditing institutions can move upstream in AI governance: not just auditing outcomes, but auditing the systems that produce them. This meta-auditing function will likely grow in importance as algorithmic systems become embedded in everything from benefit determinations to regulatory enforcement.

These cases illustrate a spectrum of on the one hand, national audit bodies such as the Netherlands Court of Audit and the U.S. GAO are beginning to define algorithmic accountability as part of their mandate. On the other, municipal cases like São Paulo reveal the dangers of conflating predictive correlation with legal culpability. Meanwhile, Estonia shows how even in highly digitized contexts, explainability and human comprehension remain critical bottlenecks.

Taken together, these cases affirm several key insights:

- AI does not eliminate the need for human judgment; it repositions it.
- Transparency is not automatically produced by digitization; it must be designed and institutionally embedded.
- Audit institutions must develop new capabilities—not only to use AI but to audit it meaningfully.

Ultimately, these examples underline that the audit of algorithms is itself a form of governance, requiring ethical reflection, procedural safeguards, and democratic legitimacy. Without these, the use of AI in public oversight may deepen the very injustices it seeks to detect.

5. IMPLICATIONS: RETHINKING ACCOUNTABILITY, ETHICS, AND DEMOCRATIC OVERSIGHT

The increasing integration of AI into the audit functions of public administration demands a fundamental rethinking of how we understand accountability, ethics, and democratic oversight in the digital age. The transition from manual, rule-based audit procedures to AI-enhanced, data-driven systems is not merely a matter of administrative modernization. It represents a structural transformation in the relationships among humans, machines, and institutions. At its core, this shift compels us to interrogate the ontological status of AI in auditing: "Is AI a tool, a collaborator, or an autonomous agent?".

The conventional framing of AI in governance positions it as a neutral instrument—a tool that enhances human capacity by processing large datasets, detecting anomalies, and improving efficiency. This instrumentalist view holds that ultimate responsibility remains with human decision-makers, while AI serves merely to support or augment decision processes (Morley et al., 2021: 3–5). From this perspective, AI is not an agent but a sophisticated calculator, constrained by its programming and incapable of moral or political judgment. However, this framing becomes increasingly inadequate in light of the autonomous learning capabilities and systemic influence of contemporary AI systems. In many public audit settings, AI systems not only identify risks or irregularities but also generate rankings, prioritize cases, and recommend actions that are rarely questioned by human actors. When auditors defer to algorithmic outputs without sufficient interrogation, AI begins to function less as a tool and more as a de facto agent—one that shapes administrative behavior and decision-making through its internal logic and perceived authority (Pasquale, 2015: 9).

This emergent agency is not intentional in the philosophical sense, but it is functional and institutionally embedded. AI systems, once deployed, become actors in governance systems, generating consequences, influencing priorities, and reshaping accountability chains. In such contexts, the line between tool and agent blurs, necessitating new ethical frameworks for assigning responsibility when things go wrong. The traditional audit mantra—who did what, when, and why—becomes difficult to apply when decisions emerge from distributed interactions between data, models, and human oversight (Wieringa, 2020: 396–398).

The concept of algorithmic discretion has been examined by scholars in different ways. For example, Bovens (2007: 451–452) emphasized human accountability within bureaucratic discretion, whereas Wieringa (2020: 395–398) and Yeung (2018: 505–509) highlight how algorithmic systems distribute agency across socio-technical networks. This article extends these debates by demonstrating how AI-based auditing repositions discretion as embedded in probabilistic modeling, thereby shifting responsibility from identifiable officials to opaque technical infrastructures. This comparative engagement clarifies the article's contribution: to show that auditing, as a site of public accountability, is uniquely vulnerable to the legitimacy challenges posed by algorithmic discretion. The shift toward AI-mediated auditing thus compels a reconfiguration of human–machine relationships in public governance. This reconfiguration is not simply technical but deeply ethical and political. It calls for institutions to reassert the normative foundations of public administration—fairness, transparency, and accountability—even as they embrace new technologies. A key challenge is to avoid both technocratic determinism (the belief that machines know best) and digital romanticism (the notion that participation or inclusion alone can fix structural inequalities).

One implication of this reconfiguration is the need to define new roles and responsibilities for auditors, public officials, and citizens. Auditors must become not only the users of AI but also the interpreters and curators of algorithmic insight (Logg et al., 2019). They must understand the assumptions, training data, and model limitations that underlie AI outputs, and be prepared to challenge or contextualize them. This requires a significant investment in algorithmic literacy, not just among technical staff but across the audit institution as a whole (Yeung, 2018: 507–509). Likewise, public officials must learn to navigate the new epistemic terrain of AI-informed governance. They must resist the tendency to equate algorithmic outputs with objective truth and instead ask critical questions about how knowledge is produced, what forms of bias are encoded, and whose interests are served by automated decisions. This demands a return to administrative ethics—not as an abstract code but as a lived practice of reflective judgment under conditions of uncertainty and complexity (Bovens, 2007: 451–452).

For citizens and civil society actors, the reconfiguration of human–machine relations raises both risks and opportunities. On the one hand, opaque and autonomous AI systems threaten to alienate the public from governance, creating a sense of disempowerment and mistrust. On the other hand, if properly governed, AI could enhance participatory oversight, enable more granular transparency, and strengthen citizen monitoring of public institutions. To realize this potential, however, citizens must be equipped with the tools and platforms to understand and contest algorithmic decisions. This transformation also has implications for institutional design. Existing audit frameworks are often not equipped to evaluate AI systems, particularly in terms of explainability, data provenance, and ethical compliance. New audit modalities—such as algorithmic impact assessments, ethics review boards, and interdisciplinary audit teams—are needed to supplement traditional approaches. These modalities must be embedded in law, supported by resources, and linked to meaningful enforcement mechanisms.

Moreover, democratic oversight must extend beyond individual systems to the ecosystems in which AI is developed, deployed, and governed. Procurement practices, vendor accountability, public-private partnerships, and software licensing arrangements all shape the functioning of administrative machines. Without clear rules on ownership, access, and public interest safeguards, governments risk becoming dependent on proprietary logics that are neither transparent nor contestable. An emerging normative framework for governing AI in public auditing must therefore be multi-dimensional. It must combine technical transparency (e.g., documentation, explainability, audit trails); procedural accountability (e.g., due process, redress mechanisms); epistemic pluralism (e.g., incorporating diverse forms of knowledge and experience); and deliberative participation (e.g., co-design of systems, participatory evaluation). At stake is not only the effectiveness of audits but the very nature of public authority in the digital age. If machines are to be trusted components of democratic institutions, they must be embedded in practices of legitimation, accountability, and ethical constraint. Otherwise, the audit function risks being captured by logics of efficiency that eclipse its foundational role in upholding justice and public trust.

Finally, this reconfiguration invites us to reflect on the broader ontology of governance. The rise of administrative machines signals a movement toward hybrid systems of power—neither fully human nor fully mechanical, but always relational and evolving. Governance in this context becomes a matter not only of decision-making but of socio-technical choreography, where institutions must continually recalibrate their structures, norms, and epistemologies to remain responsive and legitimate. In this light, AI should not be viewed merely as a new tool in the bureaucratic toolbox but as a constitutive force that reshapes what it means to govern, to audit, and to be accountable. The challenge is not to resist this transformation but to guide it democratically—ensuring that the future of public oversight remains anchored in the principles of justice, transparency, and the dignity of the governed.

CONCLUSION: TOWARD A NORMATIVE FRAMEWORK FOR ALGORITHMIC AUDITING

The incorporation of AI into the core practices of public auditing marks a profound transformation in the architecture of democratic governance. What was once a human-centered, document-intensive, and procedurally bounded activity is rapidly evolving into a data-driven, predictive, and technologically mediated domain of oversight. This shift—from rule-based verification to algorithmic surveillance—challenges the foundational principles that historically legitimized the audit function: transparency, accountability, and impartiality. Throughout this article, we have examined how AI not only augments audit capacity but also redefines the logics of public oversight.

Algorithmic systems introduce new modes of rationality, redistribute epistemic authority, and destabilize traditional accountability chains. In doing so, they raise pressing normative questions: "Who is responsible for automated decisions?", "How can algorithmic outputs be meaningfully challenged?" and "What does it mean to audit in a world governed increasingly by predictive machines?". Importantly, these transformations parallel and extend long-standing debates on continuous auditing and continuous monitoring (Vasarhelyi & Halper, 1991; Alles et al., 2006), demonstrating that AI represents not a break from but an intensification of prior technological trajectories. To address these questions, we must move beyond the dichotomy

of technological optimism and institutional inertia. A normative framework for algorithmic auditing must be built not on technical functionality alone but on a coherent set of democratic values that can guide the design, deployment, and evaluation of AI systems in the public sector. Four principles are especially critical: transparency, auditability, contestability, and institutional ethics.

A framework built on four principles—transparency, auditability, contestability, and institutional ethics—offers a way to reconcile technological efficiency with democratic legitimacy. This framework is designed to respond not only to classical bureaucratic concerns but also to the demands of AI-driven continuous auditing and predictive governance. A normative framework for algorithmic auditing must be built not on technical functionality alone but on a coherent set of democratic values that can guide the design, deployment, and evaluation of AI systems in the public sector. Four principles are especially critical:

- Transparency: Beyond disclosure of code or model structures, transparency entails communicative accessibility—the capacity of affected actors, from auditors to citizens, to understand the logic, assumptions, and limitations of AI systems (Ananny & Crawford, 2018: 975–976).

- Auditability: Systems must be traceable, verifiable, and documented to ensure meaningful scrutiny and prevent accountability gaps (Burrell, 2016: 1–3).

- Contestability: Individuals and communities affected by algorithmic decisions must have accessible mechanisms to challenge, appeal, or demand justification (Wieringa, 2020: 397–400; Eslami et al., 2025: 8–10).

- Institutional Ethics: Ethical reasoning must be embedded across all stages of system design and governance, ensuring that efficiency does not override justice and equity (Morley et al., 2021: 5–7).

The synthesis of these principles points toward a reconstructed audit paradigm—one in which the promise of AI is harnessed within a framework of democratic constraint. AI should not be viewed merely as a tool of administrative efficiency but as a constitutive element of governance, requiring safeguards against capture and erosion of oversight (Terzis et al., 2024: 7–9; Volodina & Grossi, 2024: 1450–1455; Aldemir & Uçma Uysal, 2025: 3–5). International

policy frameworks underscore that effective democratic oversight requires harmonized standards for algorithmic transparency across jurisdictions (GPAI/ OECD, 2024: 18). Ensuring that algorithmic systems align with democratic values requires not only technical fixes but also institutional courage, normative clarity, and civic vigilance, reinforced by participatory oversight frameworks (Eslami et al., 2025: 11–12). The challenge is not to automate auditing but to reimagine it—expanding its values, retooling its practices, and regrounding its legitimacy for a world shaped by administrative machines. With thoughtful design and democratic oversight, algorithmic auditing can become a powerful mechanism for enhancing accountability, illuminating injustice, and strengthening the social contract.

Any oversight system that emphasizes transparency must guard against superficial adoption of algorithm registers: without robust institutional design, clear scope, enforceability, and meaningful community participation, registers risk being performative rituals rather than instruments of accountability (Nieuwenhuizen, 2024: 405–407). Recent research also emphasizes that transparency should function as a 'pipeline' for building trust rather than a mere reputational prism (Park & Yoon, 2025: 6). Systematic reviews of algorithm auditing highlight persistent gaps, including one-shot audits that fail to detect biases continuously, and limited community involvement in oversight (Funda, 2025: 3–5). Finally, the experience of large audit firms adopting AI shows both the opportunities and major risks in terms of liability, ethics, and professional competence (Kokina, 2025: 2–4).

REFERENCES

Afzal, M. (2025). How public officials perceive algorithmic discretion. Public Administration Review. Advance online publication. https://doi.org/10.1111/puar.13957

Aldemir, C., & Uçma Uysal, T. (2025). Artificial intelligence for financial accountability and governance in the public sector: Strategic opportunities and challenges. Administrative Sciences, 15(2), 58, 1–18. https://doi.org/10.3390/admsci15020058

Alles, M., Kogan, A., & Vasarhelyi, M. A. (2006). Continuous auditing: A new view. Auditing: A Journal of Practice & Theory, 25(2), 207–235.

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973–989.

Black, J. (2005). The emergence of risk-based regulation and the new public risk management in the United Kingdom. Public Law, 3(Autumn), 512–549.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. European Law Journal, 13(4), 447–468.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1–12.

Damar, M., Köse, H. Ö., Cagle, M. N. and Özen, A. (2024). Mapping the Digital Frontier: Bibliometric and Machine Learning Insights into Public Administration Transformation. TCA Journal/Sayıştay Dergisi, 35(132), 9-41. https://doi.org/10.52836/sayistay.1455036

Danaher, J., Hogan, M. J., Noone, C., Kennedy, R., Behan, A., & Griffith, J. (2017). Algorithmic governance: Developing a research agenda through the power of collective intelligence. Big Data & Society, 4(2), 1–21.

Decker, M. C., Wegner, L., & Leicht-Scholten, C. (2025). Procedural fairness in algorithmic decision-making: The role of public engagement. Ethics and Information Technology, 27(1), 1–15. https://doi.org/10.1007/s10676-024-09811-4

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

Eslami, M., Kumar, S., & Sandvig, C. (2025). From margins to the table: Charting the potential for public participation in public-sector ADM. In Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (pp. 1–13). ACM. https://doi.org/10.1145/3715275.3732173

Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

Funda, V. (2025). A systematic review of algorithm auditing processes to assess bias and risks in AI systems. Journal of Informatics and Policy Design, 10(1), 1–18. https://systems.enpress-publisher.com/index.php/jipd/article/view/11489

GAO (U.S. Government Accountability Office). (2021). Artificial intelligence: An accountability framework for federal agencies and other entities (GAO-21-519SP). https://www.gao.gov/products/gao-21-519sp

GPAI/OECD (2024). Algorithmic transparency in the public sector: A state-of-the-art report of transparency instruments. OECD. https://wp.oecd.ai/app/uploads/2024/12/14-Algorithmic-Transparency-in-the-Public-Sector-A-state-of-the-art-report-of-transparency-instruments.pdf

Kokina, J. (2025). Challenges and opportunities for artificial intelligence in auditing by large public accounting firms. International Journal of Accounting Information Systems, 56, 100671. https://doi.org/10.1016/j.accinf.2025.100671

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103.

Mahroof, K., Weerakkody, V., Hussain, Z., Irani, Z., & Dwivedi, Y. K. (2025). Navigating power dynamics in the public sector through AI-driven algorithmic decision-making. Government Information Quarterly, 42(3), 101923, 1–12. https://doi.org/10.1016/j.giq.2025.101923

Mökander, J., Axente, M., & Floridi, L. (2021). Confronting ethical challenges in AI. AI & Society, 36(1), 59–67.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Science and Engineering Ethics, 27(4), 1–22.

Murikah, W. (2024). Bias and ethics of AI systems applied in auditing. Journal of Accounting and Public Policy, 43(4), 100118, 1–12. https://doi.org/10.1016/j.jaccpubpol.2024.100118

Netherlands Court of Audit. (2021). Use of algorithms in central government: Risks and opportunities. https://english.rekenkamer.nl/publications

Nieuwenhuizen, E. (2024). Algorithm registers: A box-ticking exercise or meaningful tool for transparency? Information Polity, 29(4), 397–413. https://doi.org/10.1177/15701255241297107

OECD (2019). Auditing in the digital era: Opportunities, challenges and implications. OECD Publishing.

OECD (2022). The future of auditing in the digital era: Emerging technologies and governance challenges. OECD Publishing.

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown Publishing.

Park, S., & Yoon, J. (2025). AI algorithm transparency, pipelines for trust not prisms: Mitigating general negative attitudes and enhancing trust toward AI. Humanities and Social Sciences Communications, 12, 515. https://doi.org/10.1057/s41599-025-05116-z

Parviainen, J., Koski, A., Eilola, L., Palukka, H., Alanen, P., & Lindholm, C. (2025). Building and eroding the citizen-state relationship in the era of algorithmic decision-making: Towards a new conceptual model of institutional trust. Social Sciences, 14(3), 178, 1-17. https://doi.org/10.3390/socsci14030178

Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.

Pérez-Durán, I. (2024). Twenty-five years of accountability research in public administration: Authorship, themes, methods, and future trends. International Review of Administrative Sciences, 90(4), 1161–1188. https://doi.org/10.1177/00208523231211751

Polat, M. (2024). Yapay Zekanın Denetimde Kullanılması ve Etik Sorunlar. Sayıştay Dergisi, 35(134), 395-423. https://doi.org/10.52836/sayistay.1554497

Power, M. (1997). The audit society: Rituals of verification. Oxford University Press.

Strathern, M. (1997). 'Improving ratings': Audit in the British university system. European Review, 5(3), 305–321.

Terzis, P., Veale, M., & Gaumann, N. (2024). Law and the emerging political economy of algorithmic audits. In Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (pp. 1–12). ACM. https://doi.org/10.1145/3630106.3658970

Vasarhelyi, M. A., & Halper, F. B. (1991). The continuous audit of online systems. Auditing: A Journal of Practice & Theory, 10(1), 110-125.

Volodina, T., & Grossi, G. (2024). Digital transformation in public sector auditing: Between hope and fear. Public Management Review, 27(5), 1444-1468. https://doi.org/10.1080/14719037.2024.2402346

Yavuz, E. (2025). Digitalization in the Fight Against Corruption: Feasibility of a Blockchain-Based System in Türkiye. TCA Journal/Sayıştay Dergisi, 36 (137), 255-287. DOI: https://doi.org/10.52836/sayistay.1675048

Yener, M., Charoenpol, M., Suntharanurak, S. and Köse, H.Ö. (2025). Strategic Cooperation of Supreme Audit Institutions of Thailand and Türkiye for Digital Transformation and Innovation in Public Sector Auditing. TCA Journal/Sayıştay Dergisi, 36 (136), 9-34. https://doi.org/10.52836/sayistay.1633666

KAMU DENETİMİNDE ALGORİTMİK SİSTEMLER VE DEMOKRATİK DENETİM

Hamza ATEŞ

GENİŞLETİLMİŞ ÖZET

Bu makale, yapay zekânın (YZ) kamu denetiminin kavramsal, süreçsel ve kurumsal temellerini nasıl dönüştürdüğünü incelemektedir. Kamu denetimi, modern kamu yönetimi geleneklerinde uzun yıllar boyunca kamu kaynaklarının kullanımında şeffaflık, etkinlik ve hesap verebilirliği sağlamak için geliştirilmiş bir mekanizma olarak görülmüştür. Hukuka uygunluk, tarafsızlık, izlenebilirlik ve usule bağlı rasyonalite gibi ilkeler üzerine kurulan klasik denetim anlayışı, esas itibarıyla geçmişte gerçekleşen işlemlerin doğruluğunu, mevzuata uygunluğunu ve sorumlularını tespit etmeye yönelik ex-post bir faaliyetti. Ancak algoritmik yönetişimin yükselişi ile birlikte kamu denetiminin bu geleneksel rolü köklü bir dönüşümden geçmektedir. Yapay zekâ sistemlerinin yolsuzluk tespitinden performans izlemeye, risk analizinden öngörücü uyum denetimine kadar geniş bir alanda kullanılmaya başlanması, denetim kavramının kendisini yeniden düşünmeyi zorunlu kılmaktadır.

Makale, yapay zekâyı yalnızca teknik kapasiteyi artıran bir araç olarak değil; aynı zamanda denetimin epistemolojik temellerini, kamu otoritesinin dağılımını ve kamusal gözetimin etik sınırlarını yeniden belirleyen dönüştürücü bir güç olarak konumlandırmaktadır. Zira YZ, karar alma süreçlerini veriye dayalı olasılıksal modellere, örüntü tanıma mekanizmalarına ve öngörülebilir risk matrislerine kaydırmakta; böylece hesap verebilirliği geçmişe dönük doğrulamadan geleceğe dönük ihtiyat ve risk yönetimine yöneltmektedir. Bu durum, kural bağlılığı ile yönetsel takdir arasındaki klasik ilişkiyi dönüştürdüğü gibi, algoritmik sistemlere duyulan artan kurumsal güven—özellikle de şeffaf olmayan "kara kutu" modeller söz konusu olduğunda—insan yargısı ile makine çıkarımı arasındaki sınırın bulanıklaşmasına yol açmaktadır. Böylece dağılmış sorumluluk, zayıflayan izlenebilirlik ve meşruiyet tartışmaları kamu denetiminin merkezine yerleşmektedir.

Makale ayrıca YZ'yi 1990'lardan itibaren geliştirilen sürekli denetim (continuous auditing – CA) tartışmalarıyla ilişkilendirmektedir. Sürekli denetim, dijitalleşme sayesinde finansal ve operasyonel verilerin anlık olarak izlenmesini ve düzensizliklerin hızla tespit edilmesini amaçlamıştı. CA, yüksek veri işleme

kapasitesi ve gerçek zamanlı izleme imkânı sunarak denetim süreçlerini modernize etmeyi vaat ediyordu. Bununla birlikte CA, yine de açık, kural-temelli, doğrulanabilir süreçlere dayanıyordu. Yapay zekâ tabanlı denetim ise bu çizgiyi daha da ileri taşıyarak öğrenen, uyarlanabilir ve çoğu zaman şeffaflığı sınırlı olan modellerle çalışmakta; böylece açıklanabilirlik (explainability), adalet (fairness) ve usule uygunluk (due process) konusunda daha derin kurumsal sorular ortaya çıkarmaktadır. Dolayısıyla YZ denetimi, sürekli denetim yaklaşımıyla hem süreklilik hem de kopuş ilişkisi içinde olup, güçlü yönlerinin yanı sıra yeni türden yönetişim riskleri yaratmaktadır.

Makalenin analiz kısmı, farklı yönetişim bağlamlarında yürütülen dört ampirik örnek üzerinden yapılandırılmıştır. Hollanda çocuk yardımı skandalı, algoritmik risk puanlamasının zayıf insan denetimiyle birleştiğinde sistematik ayrımcılığa ve ağır toplumsal zararlara yol açabileceğini dramatik biçimde göstermektedir. Estonya'nın kamu ihale denetiminde öngörücü analitik kullanımına yönelik modeli, yüksek güven kültürüne sahip dijital devletlerde dahi yorumlanabilirlik ve algoritmik önyargı risklerinin devam ettiğini ortaya koymaktadır. São Paulo belediyesinin otomatik anomali tespit sistemi ise verimlilik ile usule ilişkin güvenceler arasında yaşanan gerilimi göstermekte; vatandaşlara yönelik yaptırımların yeterli soruşturma yapılmadan uygulanabildiğini kanıtlamaktadır. Buna karşılık ABD Sayıştayı (GAO), etik çerçeveler, algoritmik denetim standartları ve kurumsal rehberler geliştirerek teknolojik dönüşüme proaktif bir kurumsal uyum örneği sunmaktadır. Uluslararası kuruluşların son dönemdeki politika belgeleri de dünya genelinde denetim kurumlarının bu dönüşüme yanıt verme zorunluluğunu vurgulamaktadır.

Bu değerlendirmelerden hareketle makale, yapay zekânın denetim süreçlerini verimlilik, hız, risk öngörüsü ve geniş veri işleme kapasitesi bakımından güçlendirebileceğini; ancak aynı zamanda önyargı, şeffaflık eksikliği, yetki yoğunlaşması ve usuli güvencelerin aşınması gibi kırılganlıkları artırabileceğini savunmaktadır. Tartışmanın odağı "YZ denetimde kullanılmalı mı?" sorusu değil, "Hangi normatif çerçeve altında kullanılmalı?" sorusudur. Bu nedenle makale, dönüşümün yönetilebilmesi için dört ilkeden oluşan bir çerçeve önermektedir: Şeffaflık (karar süreçlerinin anlaşılabilir ve açıklanabilir olması), denetlenebilirlik (sistemlerin bağımsız doğrulama ve izlenebilirlik ilkelerine uygun tasarlanması), itiraz edilebilirlik (bireylerin algoritmik kararlara karşı etkili başvuru yollarına sahip olması), ve kurumsal etik (sistem tasarımı ve kullanımında kamu yararı, eşitlik ve temel hakların merkezi bir değer olarak korunması).

Makalenin temel iddiası, algoritmik denetimin yalnızca teknolojik bir modernizasyon hamlesi değil; kamu yönetişimi açısından anayasal nitelikte bir yeniden yapılanma aracı olması gerektiğidir. YZ, otoritenin dağılımını, kurumsal takdir alanlarını ve demokratik meşruiyet mekanizmalarını yeniden şekillendirirken; kurumlar hem teknolojik determinizmi hem de eleştirisiz teknolojik iyimserliği reddetmeli, denetimi daha katılımcı, refleksif, hesap verebilir ve etik olarak temellendirilmiş bir kamusal pratik olarak yeniden tasarlamalıdır. Böyle bir yaklaşım ile yapay zekâ, demokratik kurumların şeffaflığını ve hesap verebilirliğini zayıflatmak yerine güçlendiren bir araç hâline gelebilir.