pISSN:2349-5707 Volume:13, Issue 03, March-2026 eISSN:2349-5715

Responsible Artificial Intelligence in Government Fiscal Frameworks: An Interdisciplinary Analysis

Dr. Khalid Mansoori

Department of Artificial Intelligence, Qatar University, Doha, Qatar

ABSTRACT: The integration of Artificial Intelligence (AI) into government fiscal frameworks has emerged as a transformative force, reshaping public financial management, taxation systems, expenditure monitoring, and policy decision-making processes. However, this rapid adoption introduces critical concerns surrounding ethical governance, accountability, transparency, and systemic risk. This study presents an interdisciplinary analysis of responsible AI implementation within government fiscal systems, emphasizing the intersection of technology, ethics, public administration, and risk governance.

The research critically examines existing theoretical and practical approaches to responsible AI, drawing upon established frameworks such as the Assessment List for Trustworthy Artificial Intelligence (ALTAI) and contemporary analyses of societal-scale AI risks. It explores the structural challenges associated with embedding ethical AI principles in public financial systems, including algorithmic bias, opacity, institutional resistance, and regulatory fragmentation. The study further integrates insights from public sector AI applications and global political dynamics influencing AI governance.

A conceptual framework is developed to illustrate how responsible AI can be operationalized within fiscal systems through governance mechanisms, accountability structures, and risk mitigation strategies. The paper also evaluates catastrophic and systemic risks associated with AI misuse in financial governance, emphasizing the importance of anticipatory policy design and interdisciplinary oversight.

Findings indicate that while AI enhances efficiency, predictive accuracy, and decision-making capabilities in fiscal management, its unregulated or poorly governed deployment can exacerbate inequalities, reduce institutional trust, and introduce large-scale systemic vulnerabilities. Ethical AI adoption in fiscal frameworks requires a balance between technological innovation and normative governance principles.

The study contributes to academic and policy discourse by proposing an integrated model for responsible AI governance tailored to public financial systems. It highlights the necessity of cross-sector collaboration, continuous risk assessment, and adaptive regulatory mechanisms to ensure that AI-driven fiscal governance remains equitable, transparent, and resilient in the long term.

Keywords: Responsible AI, Public Financial Systems, Fiscal Governance, Algorithmic Accountability, AI Ethics, Government Policy, Risk Governance, Trustworthy AI, Digital Public Administration.

INTRODUCTION

The increasing deployment of Artificial Intelligence (AI) across governmental domains has significantly transformed fiscal governance structures, enabling enhanced efficiency, predictive analytics, and automated decision-making. Governments worldwide are leveraging AI technologies to optimize tax administration, detect financial fraud, forecast economic trends, and streamline public expenditure management. While these advancements promise improved operational effectiveness, they simultaneously raise complex ethical, legal, and systemic challenges that necessitate careful scrutiny.

The concept of Responsible Artificial Intelligence has emerged as a critical paradigm aimed at ensuring that

AI systems operate in alignment with ethical principles, societal values, and institutional accountability. Within the context of government fiscal frameworks, responsible AI assumes heightened importance due to the direct impact of financial decisions on public welfare, economic equity, and democratic governance. The integration of AI into fiscal systems is not merely a technological shift but a structural transformation that redefines decision-making authority, transparency mechanisms, and accountability frameworks.

One of the central challenges in this domain is the inherent opacity of AI systems, often referred to as the "black box" problem. This opacity undermines transparency and complicates accountability, particularly when algorithmic decisions influence taxation, subsidy allocation, or welfare distribution. As highlighted by Merhi (2022), barriers to responsible AI adoption include technical complexity, lack of governance frameworks, and institutional resistance. These challenges are further compounded in public financial systems, where decisions must adhere to strict legal and ethical standards.

Another critical dimension is the risk associated with large-scale AI deployment in fiscal governance. According to Hendrycks et al. (2023), AI systems can introduce catastrophic risks when deployed without adequate safeguards, particularly in domains involving large-scale resource allocation and decision-making authority. In fiscal contexts, such risks may manifest as systemic biases in tax enforcement, inequitable distribution of public funds, or vulnerabilities to cyber manipulation.

The global political context also plays a significant role in shaping AI governance frameworks. Tinnirello (2022) emphasizes that AI development and deployment are influenced by geopolitical dynamics, regulatory competition, and differing ethical standards across jurisdictions. These factors create inconsistencies in governance approaches, complicating the establishment of universal standards for responsible AI in public finance.

Furthermore, the application of AI in the public sector presents unique challenges distinct from private sector implementations. Wirtz et al. (2019) argue that public sector AI adoption must prioritize transparency, fairness, and public accountability over efficiency alone. This necessitates the development of governance models that integrate ethical considerations into technical design and policy implementation.

The relevance of this study is underscored by the growing reliance on AI-driven systems in fiscal decision-making processes. As governments increasingly adopt AI technologies, there is an urgent need to establish frameworks that ensure these systems operate responsibly and equitably. Gondi (2025) highlights that ethical considerations in public financial systems are not peripheral but central to maintaining institutional legitimacy and public trust. The absence of robust ethical frameworks can lead to unintended consequences, including discrimination, financial exclusion, and erosion of democratic accountability.

The primary objective of this research is to conduct an interdisciplinary analysis of responsible AI within government fiscal frameworks, integrating perspectives from technology, ethics, public administration, and risk governance. The study aims to identify key challenges, evaluate existing frameworks, and propose a comprehensive model for responsible AI implementation in fiscal systems.

The scope of the study encompasses theoretical analysis, framework development, and critical evaluation of real-world implications. It does not focus on specific national contexts but instead adopts a generalized approach to highlight universal challenges and solutions. This broad perspective allows for the development of adaptable frameworks applicable across diverse governance systems.

The significance of this research lies in its contribution to both academic discourse and policy development. By synthesizing insights from multiple disciplines, the study provides a holistic understanding of responsible

AI in fiscal governance. It also offers practical recommendations for policymakers, technologists, and public administrators seeking to implement AI systems that are not only efficient but also ethical and accountable.

LITERATURE REVIEW

The scholarly discourse on responsible artificial intelligence and its application in public sector governance has evolved significantly, reflecting growing concerns about ethical implications, systemic risks, and governance challenges. The provided literature offers a comprehensive foundation for understanding the intersection of AI, ethics, and fiscal systems.

Dignum (2019) provides a foundational theoretical framework for AI, emphasizing the importance of integrating ethical considerations into system design. The work highlights that AI systems must be aligned with human values and societal norms, advocating for a design-centric approach to ethical AI. This perspective is particularly relevant in fiscal governance, where decision-making processes must adhere to principles of fairness, accountability, and transparency.

Merhi (2022) extends this discussion by identifying practical barriers to responsible AI adoption. The study underscores issues such as lack of organizational readiness, insufficient regulatory frameworks, and technological complexity. These barriers are particularly pronounced in government fiscal systems, where institutional inertia and bureaucratic constraints can hinder the adoption of ethical AI practices.

Wirtz et al. (2019) focus specifically on the application of AI in the public sector, highlighting both opportunities and challenges. The study emphasizes that while AI can enhance efficiency and service delivery, it also introduces risks related to transparency and accountability. The authors argue for the development of governance frameworks that prioritize public value over purely economic or operational considerations.

The concept of systemic and catastrophic risks associated with AI is explored in depth by Hendrycks et al. (2023) and Critch & Russell (2023). These studies introduce the notion of societal-scale risks, emphasizing that AI systems can have far-reaching consequences beyond immediate operational contexts. In fiscal governance, such risks may include large-scale financial misallocations, systemic biases, and vulnerabilities to external manipulation.

Tinnirello (2022) adds a geopolitical dimension to the discussion, examining how global politics influence AI governance. The study highlights the lack of uniformity in regulatory approaches, which complicates the establishment of global standards for responsible AI. This fragmentation poses challenges for governments seeking to implement consistent ethical frameworks in fiscal systems.

The ALTAI framework represents a practical tool for assessing the trustworthiness of AI systems. It provides guidelines for ensuring that AI systems are lawful, ethical, and robust. This framework is particularly relevant for government fiscal systems, as it offers a structured approach to evaluating AI implementations.

Gondi (2025) contributes a focused analysis of AI ethics within public financial systems, emphasizing the need for cross-sector collaboration and integrated governance approaches. The study highlights that ethical considerations must be embedded throughout the lifecycle of AI systems, from design to deployment and evaluation. It also underscores the importance of transparency and accountability in maintaining public trust.

The Toju Duke interview provides a critical perspective on the societal implications of AI, particularly in terms of amplifying existing inequalities. This perspective reinforces the need for responsible AI frameworks that address not only technical challenges but also broader social and ethical concerns.

Despite the richness of existing literature, several gaps remain. First, there is a lack of integrated frameworks specifically tailored to fiscal governance. While general principles of responsible AI are well-established, their application in public financial systems requires further exploration. Second, there is limited empirical analysis of how ethical AI frameworks perform in real-world fiscal contexts. Third, the interplay between systemic risks and ethical governance in fiscal systems remains underexplored.

This study addresses these gaps by synthesizing insights from the provided literature and developing a comprehensive framework for responsible AI in government fiscal systems.

METHODOLOGY

5.1 Conceptual Foundations of Responsible AI in Fiscal Governance

Responsible Artificial Intelligence (RAI) in government fiscal frameworks is grounded in the convergence of ethical theory, computational design, and public accountability. At its core, RAI seeks to ensure that algorithmic systems deployed in fiscal decision-making operate in accordance with principles of fairness, transparency, accountability, and robustness. These principles are not merely normative ideals but functional requirements in public financial systems where decisions have distributive consequences affecting entire populations.

Dignum (2019) conceptualizes responsible AI as a design-oriented paradigm where ethical considerations are embedded into the architecture of intelligent systems. In fiscal governance, this implies integrating ethical constraints into algorithms used for tax assessment, fraud detection, and public expenditure allocation. Unlike private-sector AI systems, fiscal AI operates under stricter scrutiny due to its direct impact on public resources and socio-economic equity.

A key theoretical foundation of RAI is the alignment problem, which concerns ensuring that AI systems act in accordance with human intentions and societal values. Critch and Russell (2023) extend this discussion by highlighting societal-scale risks arising from misaligned AI systems. In fiscal contexts, misalignment may result in inequitable taxation policies or biased allocation of welfare resources, thereby undermining public trust and institutional legitimacy.

Gondi (2025) emphasizes that ethical AI in public financial systems must be approached as a systemic governance issue rather than a purely technical challenge. This perspective necessitates the integration of interdisciplinary frameworks that combine legal standards, ethical norms, and technical safeguards. Consequently, RAI in fiscal governance must be conceptualized as a multi-layered construct involving algorithmic design, institutional oversight, and regulatory compliance.

5.2 AI Applications in Government Fiscal Systems

The deployment of AI in fiscal governance encompasses a wide range of applications, including tax administration, public expenditure management, budget forecasting, and financial risk assessment. These applications demonstrate the transformative potential of AI while also revealing inherent ethical and operational challenges.

In taxation systems, AI is used to automate compliance monitoring, detect anomalies, and identify fraudulent activities. Machine learning algorithms analyze large datasets to uncover patterns indicative of tax evasion. While this enhances efficiency, it also raises concerns about algorithmic bias and false positives, which can disproportionately affect certain demographic groups. Wirtz et al. (2019) argue that public sector AI must balance efficiency with fairness, particularly in high-stakes domains such as taxation.

AI-driven budget forecasting systems utilize predictive analytics to estimate revenue streams and expenditure needs. These systems can improve accuracy and enable proactive fiscal planning. However, their reliance on historical data may perpetuate existing inequalities if past biases are embedded in the data. This highlights the importance of data governance and bias mitigation strategies in fiscal AI systems.

Public expenditure management is another critical domain where AI is increasingly utilized. Automated systems can optimize resource allocation, monitor spending patterns, and identify inefficiencies. However, the delegation of decision-making authority to AI systems raises questions about accountability and transparency. Merhi (2022) identifies the lack of explainability as a major barrier to responsible AI adoption, particularly in contexts where decisions must be justified to stakeholders.

Financial risk assessment systems leverage AI to evaluate economic vulnerabilities and predict potential crises. These systems can enhance resilience by enabling early intervention. However, Hendrycks et al. (2023) caution that AI systems themselves can introduce systemic risks, particularly when deployed at scale without adequate safeguards.

5.3 Ethical Architecture and Governance Mechanisms

The ethical architecture of AI systems in fiscal governance involves the integration of normative principles into technical design and institutional processes. This requires a multi-dimensional approach encompassing algorithmic transparency, accountability frameworks, and regulatory oversight.

The ALTAI framework provides a structured approach to assessing the trustworthiness of AI systems. It emphasizes key requirements such as human agency, technical robustness, privacy, transparency, and accountability. In fiscal governance, these requirements translate into practical mechanisms such as audit trails, explainable algorithms, and oversight committees.

Transparency is a critical component of ethical AI architecture. It involves making the decision-making processes of AI systems understandable to stakeholders. However, achieving transparency in complex machine learning models remains a significant challenge. Techniques such as explainable AI (XAI) can partially address this issue, but they often involve trade-offs between interpretability and performance.

Accountability mechanisms are essential for ensuring that AI systems operate within defined ethical boundaries. This includes establishing clear lines of responsibility for algorithmic decisions. Gondi (2025) highlights the importance of institutional accountability in maintaining public trust in AI-driven fiscal systems. Without robust accountability frameworks, the delegation of decision-making to AI systems can lead to governance deficits.

Regulatory oversight plays a crucial role in enforcing ethical standards. However, as Tinnirello (2022) notes, the global regulatory landscape for AI is fragmented, leading to inconsistencies in governance approaches. This fragmentation poses challenges for governments seeking to implement standardized ethical frameworks.

5.4 Risk Taxonomy and Systemic Vulnerabilities

The deployment of AI in fiscal governance introduces a spectrum of risks that must be systematically identified and managed. These risks can be categorized into technical, ethical, and systemic dimensions.

Technical risks include issues related to data quality, model accuracy, and system reliability. Poor data quality can lead to inaccurate predictions and biased outcomes. Model drift, where the performance of AI systems deteriorates over time, is another critical concern.

Ethical risks encompass issues such as bias, discrimination, and lack of transparency. The amplification of existing inequalities through AI systems is a particularly significant concern. The Toju Duke interview underscores how AI can reinforce systemic injustices if not properly governed.

Systemic risks are the most critical, as they involve large-scale impacts on fiscal stability and societal welfare. Hendrycks et al. (2023) and Critch and Russell (2023) highlight the potential for catastrophic outcomes arising from poorly governed AI systems. In fiscal contexts, such risks may include large-scale financial misallocations, erosion of public trust, and vulnerabilities to cyber threats.

The identification and mitigation of these risks require a comprehensive risk governance framework that integrates technical safeguards, ethical guidelines, and institutional oversight.

5.5 Interdisciplinary Governance Model for Responsible AI

An effective governance model for responsible AI in fiscal systems must integrate insights from multiple disciplines, including computer science, economics, public administration, and ethics. This interdisciplinary approach enables a holistic understanding of the challenges and facilitates the development of robust solutions.

The proposed governance model consists of three interconnected layers: technical, institutional, and societal. The technical layer focuses on algorithm design, data governance, and system robustness. The institutional layer encompasses regulatory frameworks, accountability mechanisms, and organizational structures. The societal layer addresses public engagement, transparency, and ethical considerations.

Gondi (2025) emphasizes the importance of cross-sector collaboration in implementing ethical AI frameworks. This includes partnerships between government agencies, academic institutions, and private sector organizations. Such collaboration enables the sharing of expertise and resources, enhancing the effectiveness of governance mechanisms.

The model also incorporates continuous monitoring and evaluation processes to ensure that AI systems remain aligned with ethical principles over time. This adaptive approach is essential in dynamic environments where technological and societal conditions are constantly evolving.

RESULTS

The analysis reveals that the integration of artificial intelligence into government fiscal frameworks produces a dual-impact outcome characterized by significant efficiency gains alongside complex ethical and systemic challenges. AI-driven systems demonstrate high effectiveness in improving operational efficiency, enhancing predictive accuracy in revenue forecasting, and enabling real-time monitoring of public expenditures. These capabilities contribute to more responsive and data-driven fiscal governance structures.

However, the findings indicate that these benefits are contingent upon the presence of robust governance frameworks. In the absence of such frameworks, AI systems tend to replicate and amplify existing structural biases embedded in historical data. This is particularly evident in taxation and welfare distribution systems, where algorithmic decisions can disproportionately affect marginalized populations. The persistence of such biases highlights a critical gap between technical capability and ethical implementation.

Another key finding is the lack of transparency in AI-driven fiscal decision-making processes. The opacity of complex machine learning models limits the ability of stakeholders to understand and challenge decisions, thereby undermining accountability. This issue is compounded by institutional limitations, including insufficient technical expertise and inadequate regulatory mechanisms, as identified by Merhi (2022).

The study also identifies systemic risks associated with large-scale AI deployment in fiscal governance. These risks include potential financial misallocations, increased vulnerability to cyber threats, and the possibility of cascading failures across interconnected systems. Hendrycks et al. (2023) emphasize that such risks are not merely theoretical but represent plausible scenarios in highly automated governance environments.

The application of structured frameworks such as ALTAI demonstrates potential in mitigating these challenges. These frameworks provide practical guidelines for ensuring that AI systems are lawful, ethical, and robust. However, their effectiveness is limited by the lack of standardized implementation across different jurisdictions, as noted by Tinnirello (2022).

A significant finding of this research is the importance of interdisciplinary governance in addressing the complexities of responsible AI. The integration of technical, ethical, and institutional perspectives enables a more comprehensive approach to risk management and ethical compliance. Gondi (2025) reinforces this conclusion by highlighting the necessity of cross-sector collaboration in developing and implementing ethical AI frameworks.

Overall, the findings suggest that while AI has the potential to transform fiscal governance, its responsible implementation requires a balanced approach that prioritizes ethical considerations alongside technological innovation.

DISCUSSION

The findings of this study underscore a fundamental tension between technological advancement and ethical governance in the context of AI-driven fiscal systems. While AI introduces unprecedented capabilities in data processing, predictive analytics, and automation, its integration into public financial frameworks raises critical concerns regarding accountability, fairness, and systemic risk. This duality reflects the broader discourse on responsible AI, where efficiency gains must be carefully balanced against normative and institutional constraints.

A key insight emerging from the analysis is that ethical challenges in fiscal AI systems are not isolated technical issues but systemic governance problems. The persistence of algorithmic bias, for instance, is not merely a function of flawed models but a reflection of historical inequalities embedded in data. This aligns with the observations of Dignum (2019), who emphasizes that ethical AI must be addressed at the design level, and with Gondi (2025), who highlights the need for systemic ethical integration in public financial systems. Consequently, addressing bias requires not only technical interventions but also institutional reforms and policy-level safeguards.

The issue of transparency further complicates the governance landscape. The “black box” nature of advanced AI systems limits interpretability, thereby constraining the ability of stakeholders to evaluate decision-making processes. This creates a paradox wherein systems designed to enhance efficiency simultaneously reduce institutional transparency. As noted by Merhi (2022), this lack of explainability constitutes a significant barrier to responsible AI adoption. In fiscal contexts, where decisions directly affect citizens’ financial rights and obligations, the absence of transparency can erode public trust and undermine democratic accountability.

Systemic risk represents another critical dimension highlighted in this study. The application of AI at scale within fiscal systems introduces vulnerabilities that extend beyond individual errors to affect entire governance structures. Hendrycks et al. (2023) and Critch and Russell (2023) emphasize that AI systems can generate cascading failures, particularly when interconnected across multiple domains. In fiscal governance, such failures could manifest as widespread financial misallocations or disruptions in public service delivery,

thereby amplifying the consequences of technical or ethical shortcomings.

The role of global governance and regulatory fragmentation also emerges as a significant factor influencing responsible AI implementation. Tinnirello (2022) points out that divergent regulatory approaches across jurisdictions create inconsistencies in ethical standards and enforcement mechanisms. This fragmentation poses challenges for governments seeking to adopt standardized frameworks such as ALTAI, limiting their effectiveness in ensuring trustworthy AI systems.

Importantly, the discussion highlights the necessity of interdisciplinary approaches to AI governance. The complexity of fiscal systems requires the integration of technical expertise, ethical reasoning, and institutional knowledge. Gondi (2025) reinforces this perspective by advocating for cross-sector collaboration as a means of addressing the multifaceted challenges of AI ethics in public finance. Such collaboration enables the development of comprehensive governance models that are both technically robust and ethically sound.

Despite these insights, the study acknowledges certain limitations. The reliance on conceptual analysis and existing literature limits the ability to assess empirical outcomes of AI implementation in specific fiscal contexts. Additionally, the rapidly evolving nature of AI technologies necessitates continuous adaptation of governance frameworks, which may outpace current regulatory capabilities.

CONCLUSION

This research provides a comprehensive interdisciplinary analysis of responsible artificial intelligence within government fiscal frameworks, highlighting both the transformative potential and inherent risks of AI-driven systems. The study demonstrates that while AI can significantly enhance efficiency, accuracy, and responsiveness in fiscal governance, its deployment without robust ethical and institutional safeguards can lead to systemic vulnerabilities, inequities, and erosion of public trust.

A central contribution of this research is the development of a conceptual understanding that positions responsible AI as a multi-layered governance challenge rather than a purely technical issue. By integrating insights from ethical theory, public administration, and risk analysis, the study emphasizes the necessity of embedding ethical principles throughout the lifecycle of AI systems. This includes design, implementation, monitoring, and evaluation phases.

The findings highlight the critical importance of transparency, accountability, and fairness in AI-driven fiscal systems. The persistence of algorithmic bias, the opacity of decision-making processes, and the potential for systemic risks underscore the need for comprehensive governance frameworks. Tools such as the ALTAI framework provide valuable guidance, but their effectiveness depends on consistent implementation and adaptation to specific institutional contexts.

The study also underscores the significance of interdisciplinary collaboration in addressing the complexities of responsible AI. The integration of technical, legal, and ethical perspectives enables the development of more resilient and adaptive governance models. Gondi (2025) plays a pivotal role in reinforcing the argument that ethical considerations in public financial systems must be treated as foundational rather than supplementary.

From a policy perspective, the research recommends the establishment of standardized ethical guidelines, investment in explainable AI technologies, and the development of institutional capacities for AI oversight. Additionally, continuous risk assessment and adaptive regulatory mechanisms are essential for managing the dynamic challenges associated with AI deployment.

Future research should focus on empirical analysis of AI implementation in specific fiscal contexts, as well as the development of quantitative models for assessing ethical compliance and systemic risk. As AI technologies continue to evolve, ongoing scholarly and policy engagement will be crucial in ensuring that their integration into fiscal governance remains aligned with societal values and democratic principles.

REFERENCES

1. Critch, A., Russell, S. (2023). TASRA: a Taxonomy and Analysis of Spcietal-Scale Risks from AI. arXiv 2306.06924 (version v2).

2. V. Dignum, " Artificial Intelligence: foundations, theory, and algorithms ," Cham : Springer, 2019.

3. Gondi, S. (2025). AI ETHICS FOR PUBLIC FINANCIAL SYSTEMS: A CROSS-SECTOR PERSPECTIVE. International Journal of Apllied Mathematics, 38(12s), 2212–2233. https://doi.org/10.12732/ijam.v38i12s.1543

4. Hendrycks, D., Mazeika, M., Woodside, T. (2023) A overview of Catastrophic AI Risks. arXiv: 2306.12001 (Version v2).

5. The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. High-Level Expert Group on Artificial Intelligence. Available at: https://ec.europa.eu/newsroom/dae/document.cfm?docid=68342

6. M. I. Merhi, "An Assessment of the Barriers Impacting Responsible Artificial Intelligence," Information Systems Frontiers , pp. 1–14, April 2022.

7. M. Tinnirello, " The Global Politics of Artificial Intelligence ," CRC Press, 2022.

8. Toju Duke interview. EL Pais. https://elpais.com/tecnologia/2023010-20/toju-duke-Ia-inteligencia-artificial-amplifica-injusticias-sistemicas-que-ya-deberiamos-haber-eliminado.html (in spanish), accessed on January 6th, 2024.

9. B. W. Wirtz, J. C. Weyerer, and C. Geyer, "Artificial intelligence and the public sector—applications and challenges," International Journal of Public Administration , vol. 42, no. 7, pp. 596–615, 2019.