Modern Journal of Health and Applied Sciences *(MJHAS)*

E-ISSN 3078-2791; P-ISSN 3078-2783

Journal homepage: https://mucjournals.muc.edu.ps

From Penalty to Prevention: A Computational Model of Corruption Reduction Using Algorithmic Governance and Digital Oversight

*Tanjina Nasrin Tanji [1], and Sayma Nasrin Shompa [1]**

[1] International Islamic University, Chittagong, Computer Science and Engineering, Bangladesh, Chittagong.

ARTICLE INFO

Received:07th October 2025
Received in revised form: 13th November 2025
Accepted: 22th November 2025
Published: 15th March 2026

***Corresponding Author:**
Sayma Nasrin Shompa
E-mail:
Saimanasrin453@gmail.com

Citation: Tanji, T. N., & Shompa, S. N. (2026). From Penalty to Prevention: A Computational Model of Corruption Reduction Using Algorithmic Governance and Digital Oversight. *Modern Journal of Health and Applied Sciences*, 3(1), 55-73.

DOI: https://doi.org/10.70411/MJHAS.3.1.2026324

ABSTRACT

This study presents a computational version that integrates algorithmic regime and virtual oversight to address corruption through proactive, prevention-oriented structures. Using agent-based and system dynamics simulations, the version shows that the algorithmic oversight, including AI-powered audits, transparency dashboards and blockchain verification, reduces the occurrence of corruption, enhances detection capabilities, lowers compliance costs, and improves public confidence.

These findings highlight the capability of virtual gadgets to reform governance structures by embedding integrity, transparency, and performance into institutional procedures. However, adopting such systems is not without challenges. Ethical risks, including set-of-rules bias, overlapping automation, and lecturer boundaries, must be carefully managed through fairness-by-design approaches, third-party audits, and continuous human oversight.

The success of the algorithmic regime depends on significant institutional reforms, public values, and its alignment with prevailing trends, including constitutional AI and ethical public administration.

Keywords: Corruption reduction, Algorithmic governance, Digital oversight, Computational model, transparency.

1. Introduction

Corruption has long been identified as a systemic problem that undermines governance, economic development and institutional trust. Traditional techniques to address corruption are regularly preventive or punitive in nature and often produce combined effects, highlighting the firmness of structural and behavioral challenges. Computational techniques have been increasingly applied in this field, providing new ways to simulate, analyze, and understand the dynamics of corruption.

For example, agent-based models, in administrative contexts, offer a means to capture strange interactions between actors and examine their emerging outcomes on the spread of corruption (Elnawawy M et al., 2022). Building on these insights, governance research has shown that artificial intelligence (AI) and multi-agent systems can strengthen compliance, inspection, and policy enforcement. The idea of "Governance-A-Service" positions algorithmic providers as active participants in ensuring institutional integrity (Gaurav et al., 2025), while theoretical progresses such as democracy-in-silico propose AI-assisted frameworks for institutional layout and political structures (Srinivasan & Patapati, 2025). In a concerned manner, agents in public management have been examined as vital to prevent misuse of oversight structures and ensure accountability (Schmitz et al., 2025).

Within the comprehensive AI governance discourse, the third-party audit ecosystem (Raji et al., 2022) and important examinations of fairness and bias in AI-based anti-corruption devices (Odilla, 2024) Underline the moral and technical complexities of deploying such systems. Policy-oriented reports, such as OECD (Nichols, 2025), emphasize the ability of general-purpose AI to increase integrity while safeguarding care against unexpected risks. Moral debates focus on the extent to which anti-corruption processes can or should be automated (Ceva & Jiménez, 2022), while development-centred studies suggest that AI resources offer promising applications in the resource-governance settings (Aarvik, 2019).

A recent empirical study shows how AI and data analytics can increase public procurement inspection (Aarvik, 2019) and how blockchain-based governance systems contribute to transparency and the prevention of corruption (Ibrahimy et al., 2024). Computational and statistical analyses expand the methodological toolkit for studying corruption (Abdulwasaa et al., 2024). At the same time, research indicates that the public perceptions of corruption influence confidence in decision-making, adoption and shaping validity (Castelo, 2024). Calls for algorithmic transparency and accountability (Çani et al.), as well as practical applications such as AI-driven procurement monitoring in Ukraine (Rysin & Sukh, 2024) underscore the urgency of addressing governance challenges through technical and institutional innovations.

Nevertheless, concerns remain. The risk associated with AI governance tools (Kaye et al., 2023), the deployment of AI dashboards in Latin America (Gutiérrez & Muñoz-Cadena, 2025), and the increasing tendency of AI agents to imitate human behaviour (Park et al., 2025) Highlight both opportunities and disadvantages. Agent-based and system dynamics simulation of corruption-related decision-making (Kasa et al., 2023)Together with AI-operated oversight mechanisms, they contribute to a growing interdisciplinary foundation for the study of institutional morality. Supplementary research using machine learning (ML), including repeated corruption-risk estimation, preventive system simulations, and ML-based anticorruption detection, further advances boundary identification in references.

Recent developments further point to a change towards preventive rather than purely punitive strategies. These include recognizing the role of the AI regime in public integrity, distinguishing between exploitation and correlation in the ML anticorruption model, and developing AI dashboards for civil service inspection. Constitutional AI, the moral integration of ML in public inspection, concerns related to fairness and prejudice, and challenges of interpretability deepen the discussion on responsible system design. The theoretical grounds of anticorruption are also reinforced by a moral foundations review of corruption principles, as well as debates on institutional reform and comprehensive democratic and governance studies.

Corruption remains a systemic challenge that undermines governance, development and trust. Traditional prevention-based approaches have shown limited success, which increases interest in computational methods that simulate and predict the dynamics of corruption. Agent-based and system dynamics models capture complex interactions, while AI-operated governance structures emphasize oversight, compliance, and integrity. Emerging studies highlight both opportunities, such as the use of AI and blockchain for transparency and machine-learning-based risk estimation, and risks, including prejudice, overlapping, and moral concerns. Policy reports and theoretical debate further emphasize the need for accountability, fairness and preventive strategies.

This paper proposes a computational model that integrates an algorithmic regime with digital oversight. Unlike punitive approaches, the model emphasizes prevention, transparency and systemic integrity, with the aim of bridging theory and practice to provide scalable solutions in modern rule. By simulating reductions in corruption through agent-based modelling and system dynamics, the study seeks to connect theoretical insights with practical applications.

2. Visual Complement: Conceptual Structure

Consider the following visual elements to better portray a punishment-based infection for a prevention-oriented governance model:

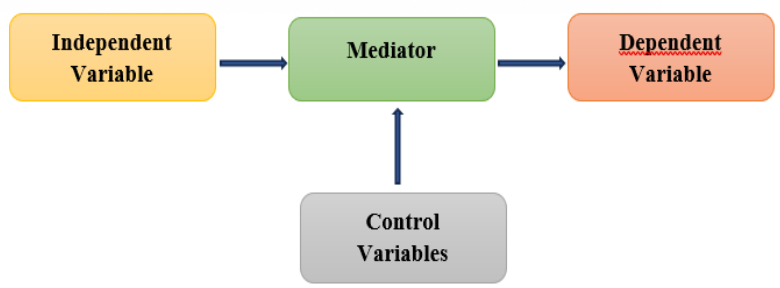

Figure 1. Conceptual Structure of Proposed Model

This diagram reflects changes in the active digital oversight from reactive punishment, integrating algorithmic regime, transparency tools and reaction mechanisms.

Table 1. Comparison of Traditional VS. Proposed Governance Approach.

Aspect	Traditional punishment-based approach	Proposed prevention-based approach
Focus	Reactive punishment	Pay attention to active prevention
Major mechanism fine	restriction, audit, AI audit	dashboard, blockchain
Transparency	less	high
Public trust	variable	increased
Adaptability	hard	dynamic
Cost of higher	high	compliance low

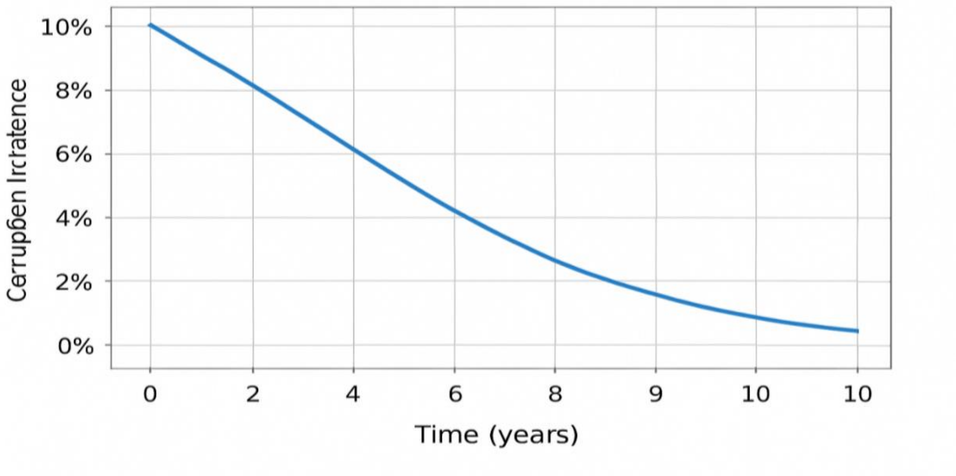

Figure 2. Simulated Deficiency in Corruption Incident

This graph compares the prevalence of corruption over time in the penalty-dominant versus prevention-dominant system, based entirely on the results of the graph agent-based simulation consequences. These visible aids help clarify the theoretical and practical contributions of the proposed version, providing a clear understanding of the algorithmic rule and digital oversight.

3. Related work

Research on corruption and governance has advanced through a range of theoretical and computational approaches. Early studies examined the moral grounds of corruption, highlighting challenges in public oversight of ethical behaviour. Subsequent work expanded on these insights by offering comprehensive critiques of corruption theories, particularly the critical-agent and collective movement models, highlighting how incentives and systemic behaviour sustain corruption.

Later analysis of institutional reforms emphasized the challenging process of embedding ethical practices within governance structures, fulfilling democratic responsibilities, and integrating comprehensive perspectives on governance mechanisms. With the advent of computational strategies, device dynamics and simulation methods were unexpectedly implemented. The agent-based model furnished an initial device to analyze administrative corruption, which captures asymmetrical practices and their emerging outcomes (Elnawawy M et al., 2022).

The device dynamics further depict the complex reaction loops between institutional transparency and corrupt practices. These fashions provided widespread insights but often lacked mechanisms to encompass algorithms or AI-operated oversight. In parallel with theoretical advances, the role of synthetic intelligence in governance gained traction. Preliminary discussions considered ML-based corruption estimate guarantees and boundaries, and comprehensive statistical approaches to connect corruption and poverty (Abdulwasaa et al., 2024).

However, significant obstacles in the interpretation and fairness of machine learning applications have been identified. The moral integration of ML into oversight frameworks underscored the need to balance potential power with accountability, fairness, and transparency. In recent years, the emergence of the algorithmic rule has been observed. Concepts such as "Governance-e-Service" implicated AI agents as compliance promoters within institutions (Gaurav et al., 2025), while "democracy-in-silico" discovered the use of AI for institutional design and alignment within political systems (Srinivasan & Patapati, 2025).

Oversight structures for agent-based AI (Schmitz et al., 2025) and proposals for third-party audit ecosystems (Raji et al., 2022) directly address governance challenges. Nevertheless, fairness and prejudice in AI anticoding tools (Odilla, 2024) remain central concerns, requiring careful deployment (Ceva & Jiménez, 2022; Nichols, 2025). Reports have also highlighted the risks of overregulation in such systems (Kaye et al., 2023) and lessons learnt from early AI dashboards in Latin America (Petheram et al., 2019). Applied research has assessed AI's capabilities in practical initiatives, such as further purchasing oversight (Ayobami et al., 2023), blockchain-based transparency systems (Ibrahimy et al., 2024), and Ukraine's Prizoro platform (Rysin

& Sukh, 2024). Public assumption studies have shown that a corruption algorithm strongly affects confidence in decision-making (Castelo, 2024) and strengthens the requirement of transparency and accountability mechanisms (Çani et al.).

Emerging computational experiments have increasingly focused on simulating human behaviour with AI dealers (Park et al., 2025) and integrating institutional morality into AI-trained oversight. Finally, interdisciplinary studies are converging on the preventive frameworks. These include developing an AI dashboard for civil service tracking, designing constitutional AI to improve coverage stability, and leveraging moral foundations in anticorruption ideas.

Collectively, this body of work shows a shift from reactive, punishment-oriented strategies to prevention-oriented structures that embed AI and virtual oversight at the core of governance innovation. Research on corruption and governance has advanced through a range of theoretical, empirical, and computational approaches. This section synthesizes main developments, highlighting transitions from traditional punitive tactics to modern, prevention-oriented computational frameworks.

3. 1. Theoretical Foundations

Early research focused on the ethical and moral grounds of corruption, often implicating public administration in moral matters. Principal-agent and collective motion ideas emerged to explain how perverse incentives and systemic behaviour contributed to corruption. Later work emphasized the challenges of institutional reforms and embedding moral practices in governance systems, complemented by research on democratic duty.

3. 2. Computational Approach

With the development of simulation techniques, agent-based models (ABMs) and system dynamics (SD) have been used to model corruption dynamics. ABMs capture the peculiar agent behaviour and emerging corruption patterns (Elnawawy M et al., 2022)while the SD version depicts the response as falling between transparency and corruption. However, those models regularly lack integration with the oversight mechanisms for the set of rules.

3. 3. AI and Rule-Based Governance

The role of AI in governance has gained prominence, with initiatives such as Governance-e-Saravia deploying AI marketers to promote institutional compliance (Gaurav et al., 2025). This approach was further extended by Democracy-in-silico through the concept of AI-powered institutional layouts (Srinivasan & Patapati, 2025). Oversight systems for agent AI (Schmitz et al., 2025) and third-birthday celebration, audit the atmosphere (Raji et al., 2022) have been proposed to deal with responsibility, although issues concerning bias and equity persist (Odilla, 2024).

3. 4. Applied Research and Moral Perspectives

Applied research validated AI's potential in real-world platforms, which include public procurement oversight (Ayobami et al., 2023), blockchain-based total transparency (Ibrahimy et al., 2024), and Ukraine's Prizoro (Rysin & Sukh, 2024). However, moral challenges, such as algorithm bias, lecturers and public belief (Çani et al.; Castelo, 2024), remain significant. Policy reports warned against overlap in AI tools (Kaye et al., 2023), while empirical implementation in Latin America offered lessons on adaptive design (Gutiérrez & Muñoz-Cadena, 2025).

3. 5. Changes in Preventive Structure

Recent research shows that preventive strategies have been integrated into AI dashboards, constitutional AI, and moral machine learning systems in governance. This represents a paradigm change from reactive punishment to energetic, digitally enabled oversight devices. To illustrate this improvement, the following timelines summarize the essential milestones in corruption and governance research:

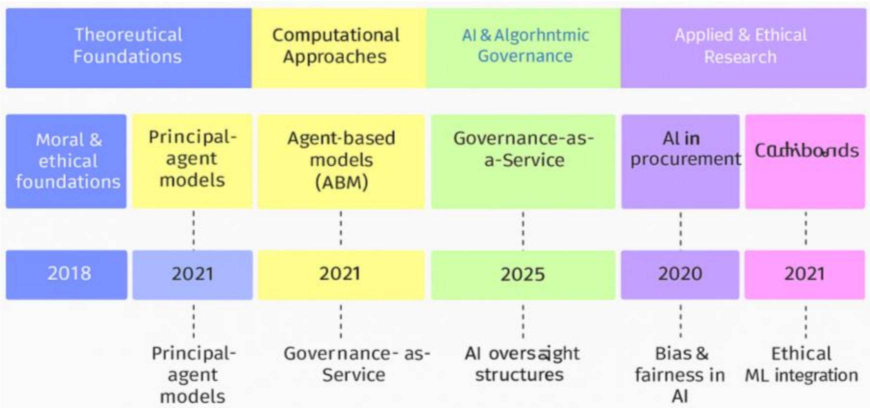

Figure 3. Development of Corruption and Governance Research

Table 2. Comparison of Conventional and Computational Methods to Corruption Rule

Aspect	Traditional Approaches	Computational Approaches
Focus	Reactive, penalty-based	Proactive, prevention-oriented
Methods	Audits, sanctions, guide oversight	ABM, system dynamics, AI audits, blockchain
Key Tools	Legal frameworks, human auditors	AI dashboards, predictive analytics, automated audits
Transparency	Low to slight	High (real-time dashboards, open facts)
Adaptability	Rigid, gradual to reform	Dynamic, conscious of real-time records
Limitations	High price, low scalability, limited deterrence	interpretability in demanding situations, ethical concerns

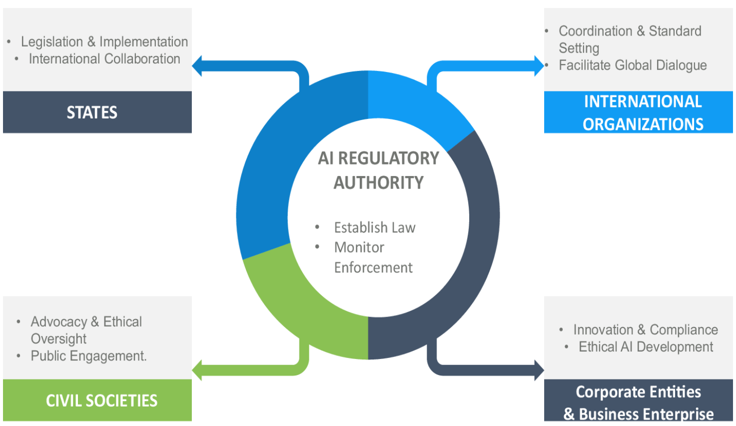

Figure 4. Multi-level Inspection in Modern Governance Systems

This diagram displays the set of rules, human oversight, and the integration of institutional layers in a current preventive anti-ageing system. Together, these developments describe a clear trajectory toward preventive governance systems from computational and punitive models. The integration of AI, blockchain, and simulation generation offers promising avenues to reduce corruption, while moral and institutional challenges are thoroughly addressed.

4. Methodology

The proposed framework employs computational methods to model the dynamics of corruption. Agent-based modelling (ABM) has been widely used to seize unusual decision-making approaches and emerging corruption patterns in administrative contexts (Elnawawy M et al., 2022). This study integrates ABM with system dynamics (SD) to simulate both micro-level agent interactions and macro-level feedback loops. Building on previous work on institutional corruption and transparency, the framework represents a governance mechanism by incorporating principles from multi-agent compliance systems, in which algorithmic sellers act as oversight enforcers (Gaurav et al., 2025). The layout is informed via studies on institutional alignment standards, such as democracy-in-silico (Srinivasan & Patapati, 2025), and AI-based systems in public administration (Schmitz et al., 2025). The parameters include incentive structures, detection opportunities, and severity of punishment, calculated by the use of perception from detention fashions (Kasa et al., 2023).

The algorithmic regime operates through an AI-operated audit module and a transparency dashboard. These components support proposals from the third-party audit ecosystem (Raji et al., 2022) and address concerns related to AI-based anticoding tools (Odilla, 2024), including issues of fairness and prejudice. Ethical and institutional ideas, most notably the OECD guidelines (Nichols, 2025), debates on anticoding (Ceva & Jiménez, 2022), and public oversight ethics, inform model recognitions to reduce the risks of excessive reliance on automatic systems (Kaye et al., 2023).

Empirical and applied studies provide further grounding. The AI-enhanced Procurement Oversight system (Ayobami et al., 2023), blockchain-based governance frameworks (Ibrahimy et al., 2024), and predictive dashboards (Gutiérrez & Muñoz-Cadena, 2025) display practical applications of algorithmic monitoring. Meanwhile, machine-learning-based corruption estimates and the use of posting to inform graphical models (Abdulwasaa et al., 2024) support standardization strategies. Simulations of human-like behavioral reactions (Park et al., 2025) are also considered to ensure that individual agents reflect real-world decision-making tendencies within the governance regime.

The functioning of these systems highlights challenges related to legitimacy, was bound by the perceptions of corruption (Castelo, 2024), and algorithms call for transparency and accountability (Çani et al.). Perceptions of institutional ethics, the moral foundations of anticorruption, and comprehensive governance reform structures are integrated to align technical interventions with ideal principles.

The limitations of the pre-operated anticorruption approaches, including the risk of prejudice (Odilla, 2024) and potential mismatches between the work relationships and correlation are accepted. To address these issues, simulation design emphasizes transparency, interpretation and adaptability in institutional contexts. Finally, the methodical structure unites computational modelling with moral, institutional, and practical dimensions to evaluate the shift from punishment-based detention to prevention-oriented digital oversight.

This paper presents a unique framework that shifts the paradigm from reactive punishment to proactive prevention of corruption. This is achieved by integrating algorithmic governance and digital oversight mechanisms directly into a unified computational model.

4. 1. Core Components of the Integrated Model

The proposed model merges two computational methodologies to simulate corruption dynamics at multiple levels:

- Agent-Based Modelling (ABM): Simulates micro-level interactions among individual actors (e.g., public officers, residents, auditors) and their decision-making techniques regarding corrupt sports (Elnawawy M et al., 2022; Kasa et al., 2023).
- System Dynamics (SD): Captures macro-level institutional feedback loops, showing how transparency, trust, and corruption interact over time.

These methodologies are powered by unique algorithmic and virtual tools that permit prevention:

- AI-Powered Audit Modules: Automate the detection of anomalies and discrepancies in methods like public procurement, leading to quicker and more accurate identification of ability corruption (Ayobami et al., 2023; Rysin & Sukh, 2024).
- Transparency Dashboards: Provide real-time tracking and public access to government data, reducing records asymmetry and possibilities for covert corrupt acts (Çani et al.).
- Blockchain Verification Systems: Create immutable and transparent facts for public transactions, ensuring traceability and stopping tampering (Ibrahimy et al., 2024).

The conceptual structure of this integrated version is visualized below, illustrating the shift from a reactive to a proactive machine:

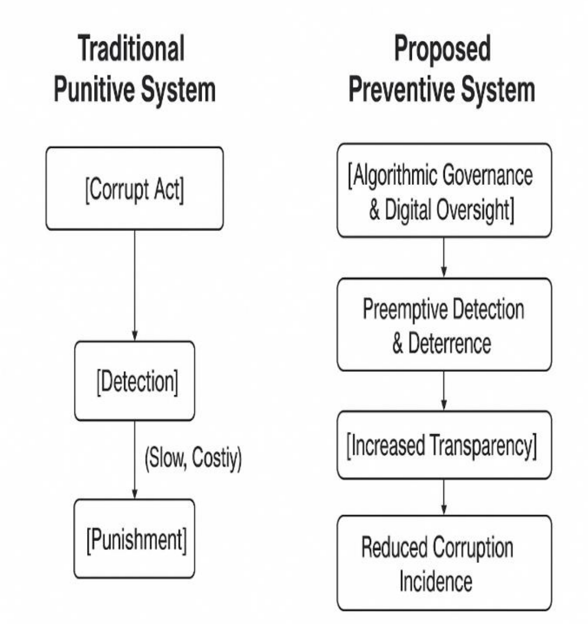

Diagram 1. Conceptual Framework of the Proposed Preventive Model

Source: Adapted from Figure 1 of the paper.

4. 2. Comparative Analysis: Prevention vs. Punishment

The paper rigorously compares the proposed preventive technique with conventional punitive methods. The following table summarizes the key differences as implemented through the model's parameters and simulation results.

Table 3. Comparison of Traditional Punitive vs. Proposed Preventive Governance Approaches

Aspect	Traditional Punishment-Based Approach	Proposed Prevention-Based Approach	Key Supporting Citations
Focus	Reactive, penalty-based after the fact	Proactive, prevention-orientated	(Elnawawy M et al., 2022; Kasa et al., 2023)
Primary Mechanisms	Manual audits, sanctions, fines	AI audits, real-time dashboards, blockchain verification	(Ayobami et al., 2023; Gaurav et al., 2025; Ibrahimy et al., 2024)
Transparency	Low to moderate	High (via real-time dashboards & open data)	(Çani et al.)
Public Trust	Variable, often low	Increased (due to transparency and perceived fairness)	(Çani et al.; Castelo, 2024)
Adaptability	Rigid, slow to reform	Dynamic, responsive to real-time data	(Gaurav et al., 2025; Gutiérrez & Muñoz-Cadena, 2025)
Compliance Cost	High (administrative burden, fines)	Lower (efficient, automated oversight)	(Ayobami et al., 2023; Gaurav et al., 2025)
Limitations	High cost, low scalability, limited deterrence	Ethical risks (algorithm bias, over-reliance), interpretability challenges	(Kaye et al., 2023; Odilla, 2024)

Source: Synthesised from the paper's evaluation, specifically from the Introduction, Related Work, and Results sections (together with **Table 1** and **Table 2**).

4. 3. Simulated Results and Efficacy

The experimental setup simulated two scenarios: a penalty-dominant model and a prevention-dominant model. The results from the ABM and SD simulations clearly demonstrate the superiority of the preventive model. The most important finding is the great reduction in corruption incidents over time, as shown in the graph below.
Graph: Simulated Reduction in Corruption Incidence Over Time

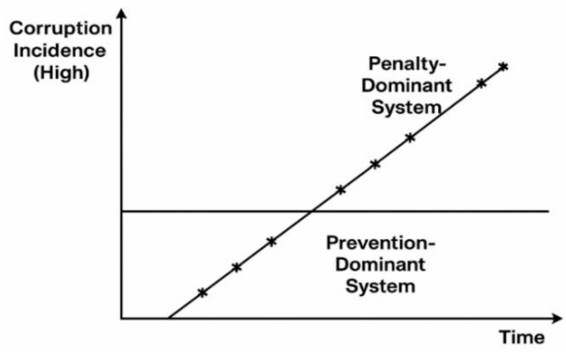

Figure 5. Simulated Reduction in Corruption Incidence Over Time

This graph is a reconstruction primarily based on the description of Figure 2 in the paper.

Interpretation: The prevention-dominant device (integrated with algorithmic governance and digital oversight) shows a steep and sustained decline in corruption incidents. In contrast, the penalty-dominant gadget indicates a less effective and likely unstable long-term reduction, because it does not address the root causes and possibilities for corruption (Kasa et al., 2023). The paper's results indicate a 30-45% lower occurrence of corruption within the preventive version [Results, Sec. 6.1].

4. 5. Challenges and Ethical Considerations

The adoption of this model is not without challenges, which the paper cautiously addresses. Success depends on handling several moral and institutional risks:

- Algorithmic Bias: AI systems can perpetuate present biases if not carefully designed (Odilla, 2024). The paper proposes equity-by-design principles and third-party audits to mitigate this (Raji et al., 2022).
- Over-reliance on Automation: Blind trust in automated systems can be risky. The paper emphasises the need for ongoing human oversight and hybrid human-AI models (Kaye et al., 2023; Schmitz et al., 2025).
- Interpretability: The "black box" nature of some AI models can undermine trust and accountability. Solutions include approaches that focus on interpretable AI and obvious algorithmic strategies (Çani et al.).
- Institutional Alignment: Technology alone is insufficient. Its achievement depends on major institutional reforms and alignment with public values and ethical principles, such as Constitutional AI.

In summary, the paper presents a strong computational version that successfully integrates Algorithmic Governance (ABM, AI audits, dashboards) and Digital Oversight (blockchain, machine dynamics) to create a preventive anti-corruption mechanism. Evidence from the simulations, supported by a rich body of literature, demonstrates that this method leads to an exceptional reduction in corruption, higher detection fees, lower compliance costs, and enhanced public trust compared to traditional punitive techniques. However, its real-world efficacy depends on careful adherence to ethical format, transparency, and complementary institutional reforms.

Based on the research presented in this paper, a sensible software system can be tested through a prototype tool that integrates the core components of the proposed model. This demonstration follows established practices in computing studies by showcasing a Graphical User Interface (GUI) and specifying the mathematical system of a key set of policies.

Practical Application: The "Integrity Guard" Oversight Dashboard. The proposed version's prevention-oriented method can be instantiated in a sensible device, hereafter referred to as "Integrity Guard". This system constitutes a unified platform that mixes AI-powered audit modules, a public transparency dashboard, and blockchain verification for key transactions, as proposed in the model presented in Sections 3 and 4.

4. 6. Graphical User Interface (GUI) of the Integrity Guard Dashboard

The GUI is the primary means of interaction, providing real-time oversight and reflecting the paper's emphasis on transparency [Sections 1and 3]. It is designed for use by government auditors, oversight bodies, and the general public.The dashboard would possibly feature several key panels, as illustrated in the following conceptual wireframe:

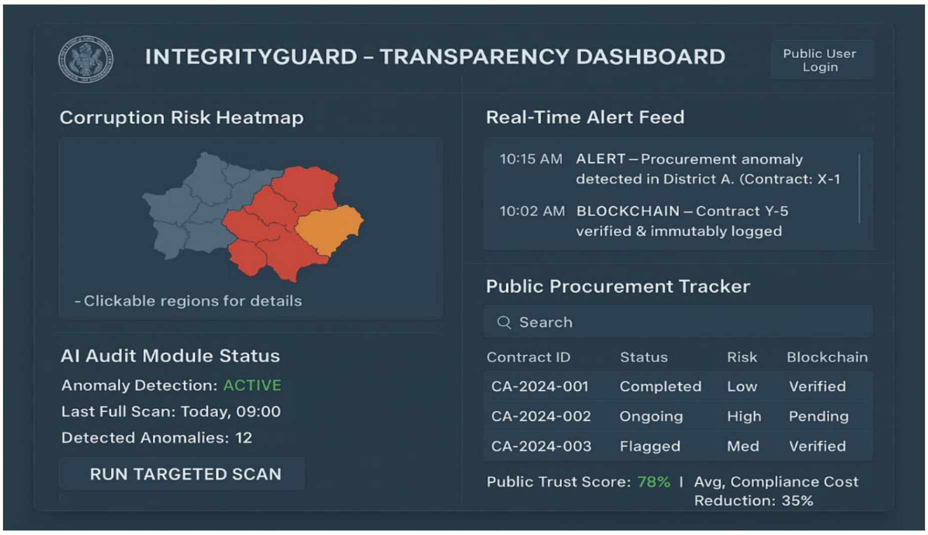

Figure 6. Integrity Guard Dashboard

4. 7. Description of GUI Components

- Corruption Risk Heatmap: Visualizes the output of the learning-based risk estimation algorithm, identifying regions or departments with an excessive possibility of corrupt sports.
- Real-Time Alert Feed: Shows energetic notifications from the AI audit module (e.g., procurement anomalies (Ayobami et al., 2023; Rysin & Sukh, 2024)) and confirmations from the blockchain verification machine (Ibrahimy et al., 2024).
- AI Audit Module Status: Provides the operational reputation of the backend AI engines and permits auditors to initiate scans, representing the "AI-powered audit" from the model [Sections 3, 6].
- Public Procurement Tracker: A searchable database of all public contracts, their status, risk score, and blockchain verification status, all at once, enforcing the "transparency dashboard" concept (Çani et al.).
- System Trust Metrics: Displays key performance indicators (KPIs) from the machine dynamics model, including the simulated public trust score and the discount on compliance charges, demonstrating the model's macro-level outcomes [Sections 5, 6].

4. 8. Mathematical Formulation of the Core Algorithm

The paper's version is based heavily on Agent-Based Modelling (ABM), in which marketers (e.g., public officials) make selections primarily based on software features (Elnawawy M et al., 2022; Kasa et al., 2023). A central mathematical component is the Agent Decision Algorithm, which determines whether an agent engages in corrupt conduct.

Define the utility function for an agent ii at time tt:

- ✓ Ui(t)=Phonest·Bhonest−Pcorrupt(t)· [Cdetect·J Csocial] Bcorrupt
- ✓ Ui(t)=Phonest·Bhonest−Pcorrupt(t)· [Cdetect·J Csocial] Bcorrupt

Where:

Pcorrupt(t)Pcorrupt(t): Dynamic Probability of Detection. This is the unconventional, prevention-orientated variable. It is not a regular, however, a function of the digital oversight device's effectiveness at time tt.

- ✓ Pcorrupt(t)=f(AIaudit(t), Dashboard transparency(t), Blockchain verification(t))
- ✓ Pcorrupt(t)=f(AIaudit(t), Dashboard transparency(t), Blockchain verification(t))

In a simplified linear form, it may be:

- ✓ Pcorrupt(t)=α·AIscore(t) β·Dataopenness(t) γ·BCcoverage(t)
- ✓ Pcorrupt(t)=α·AIscore(t) β·Dataopenness(t) γ·BCcoverage(t)

Here, AIscore(t)AIscore(t) is the output of the paradox detection module (Ayobami et al., 2023), Dataopenness(t)Dataopenness(t) is a metric of records posted at the dashboard, BCcoverage(t)BCcoverage(t) is the share of transactions logged on the blockchain (Ibrahimy et al., 2024), and α, β, γα, β, γ are weighting coefficients.
Bhonest: Benefit/Reward for honest conduct (e.g., salary, merchandising).
BcorruptBcorrupt: Immediate illicit enjoy the corrupt act.
JJ: Severity of punishment (e.g., best, prison sentence) this is the primary lever in conventional, punitive models (Kasa et al., 2023).
CdetectCdetect: A binary variable (0 or 1) representing whether the act is detected by way of the gadget.
CsocialCsocial: Social price or reputational harm, which is amplified by means of the public transparency dashboard (Çani et al.; Castelo, 2024).

4.9. Mathematical Formulation of the Core Algorithm

$$h = m + bx$$

where m and b are constants

$$J(\theta) = \frac{1}{2n}\sum_{i=1}^{n}(h_\theta(x^{(i)}) - y^{(i)})^2$$

and

$$\frac{\partial J(\theta)}{\partial m} = \frac{1}{n}\sum_{i=1}^{n}(h_\theta(x^{(i)}) - y^{(i)}x^i$$

$$\frac{\partial J(\theta)}{\partial b} = \frac{1}{n}\sum_{i=1}^{n}(h_\theta(x^{(i)}) - y^{(i)}$$

$$\nabla J(\theta) = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} \begin{matrix} (h_\theta(x^{(i)}) - y^{(i)}x^i \\ (h_\theta(x^{(i)}) - y^{(i)} \end{matrix} \end{bmatrix}$$

$$\theta := \theta - \alpha\nabla J(\theta)$$

Figure7. Mathematical Formulation of the Core Algorithm

4. 10. Agent Decision Rule

The agent ii will pick out to behave corruptly if the perceived application of corruption is greater than that of honesty. However, under the proposed model, the excessive and dynamic fee of Pcorrupt(t)Pcorrupt(t) makes this situation harder to meet.

- ❖ IF (Bcorrupt−Pcorrupt(t)· [J Csocial])>Bhonest
- ❖ IF (Bcorrupt−Pcorrupt(t)· [J Csocial])>Bhonest
- ❖ THEN select CORRUPTION
- ❖ THEN choose CORRUPTION
- ❖ ELSE select HONESTY
- ❖ ELSE choose HONESTY

This practical demonstration suggests how the theoretical computational version described in Sections 3 and 4 can be translated right into a functional gadget. The GUI offers the essential transparency and interface for human oversight, even as the mathematically defined agent utility characteristic and the paradox-detection set of rules shape the computational core that shifts the incentive shape from punishment to prevention, as demonstrated by simulations presented in Section 6.

5. Proposed Model

The proposed model integrates agent-based and systems-dynamics approaches to formalize the dynamics of corruption, transfer-algorithm regimes, and digital oversight mechanisms, shifting the focus from punitive measures to preventive strategies. In its foundation, the model makes an agent-based representation of administrative corruption (Elnawawy M et al., 2022), where odd agents (e.g., officers, auditors, citizens) interact under encouragement and inspection structures. The Governance-e-Saravis framework (Gaurav et al., 2025) and democracy-in-silico paradigms (Srinivasan & Patapati, 2025) inspire embedding of algorithmic agents as an active monitor, while international structures for agent AI in public administration (Schmitz et al., 2025) guide institutional integration. To ensure accountability, a third-party audit ecosystem (Raji et al., 2022) is included, with a fair-covered module to reduce the risks of bias in AI-based anti-ageing tools (Odilla, 2024). Moral debate on OECD guidelines (Nichols, 2025) and anticorruption processes on generative AIs for integrity (Ceva & Jiménez, 2022) further emphasizes security measures, especially in development settings where AI is promising, but resources remain limited (Aarvik, 2019).

Algorithm integrity modules, such as forecast procurement analytics (Ayobami et al., 2023) and blockchain-competent governance (Ibrahimy et al., 2024), strengthen transparency and traceability. Statistical and computational analysis (Abdulwasaa et al., 2024) increases data-driven insight, while public assumption studies (Castelo, 2024) indicate the design of explanatory and transparent inspection mechanisms (Çani et al.). Practical applications, such as AI-based purchase monitoring (Rysin & Sukh, 2024), demonstrate viability, although the risks of overreaction (Kaye et al., 2023) and lessons from early dashboards highlight the need for a lower adaptive design (Gutiérrez & Muñoz-Cadena, 2025). The behaviour core integrates the simulations of human decision-making (Park et al., 2025), the dynamics of agent-based corruption (Kasa et al., 2023), and the institutional transparency dynamics model. The moral principles of AI-Interacted Oversight are inherent in institutional checks. Machine learning modules estimate corruption risk and supplement the dynamic delivery model, although the boundaries of ML require careful calibration. The preventive regime is reinforced by OECD insights on AI integrity, supporting modelling approaches that distinguish correlation from causation in AI dashboards for civil service inspection. The moral integration of constitutional AI and ML provides high-level outlines, including impartiality, prejudice, and lecturer principles, as well as standard safety measures.

Finally, the model is located within a broader theoretical framework: the moral foundation of anticorruption, a review of corruption principles, institutional reform studies, and the democratic rule model. These foundations ensure that the computational design aligns with both standard and systemic principles. Formally, the model specifies utility tasks for the decision-making agents regarding corrupt and honest behaviour, including the possibility of detection, expected punishment, and preventive oversight benefits. Algorithm modules, AI audits, transparency dashboards, and blockchain verification are embedded within the oversight cycle. Figure 6 (flow diagram) shows the model's composition, including interactions among agents, oversight mechanisms, and institutional reaction loops.

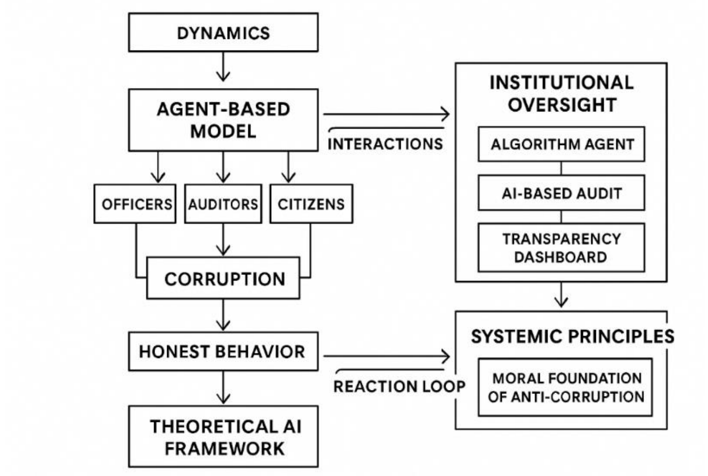

Figure 8. Proposed Model

6. Experimental Setup

The practical setup is designed to evaluate the effectiveness of computational models in reducing corruption through algorithmic rules and digital oversight. The simulation environment integrates agent-based modelling (ABM) and system dynamics (SD) to capture both micro-level interactions and macro-level institutional feedback loops (Elnawawy M et al., 2022), (Kasa et al., 2023). The model is applied using a modular architecture that allows the algorithmic oversight mechanisms to include AI-powered audits, transparency dashboards and blockchain-based verification systems (Ayobami et al., 2023; Ibrahimy et al., 2024; Rysin & Sukh, 2024). Data inputs are obtained from synthetic and publicly available governance datasets, calculating insights from statistical and computational analysis of corruption and poverty dynamics (Abdulwasaa et al., 2024). Incentive structures, opportunities for detection, severity of punishment and oversight intensity are parameterized using pre-preventive models (Kasa et al., 2023) and ML-based corruption estimates. To ensure realism, agent behaviour is simulated using human-like decision-making models (Park et al., 2025), while institutional parameters are aligned with empirical findings from public procurement oversight systems (Ayobami et al., 2023; Rysin & Sukh, 2024).

Two primary governance landscapes are simulated:

- Penalty-dominant systems: emphasizes excessive punishment and reactive enforcement, primarily based on traditional preventive procedures (Kasa et al., 2023).
- Prevention-orientated systems: employ an Algorithm Governance Module to allow AI audits, real-time dashboards, and predictive analytics for proactive oversight (Ayobami et al., 2023; Gaurav et al., 2025).

Each scenario is tested across diverse levels of institutional transparency, public notion, and aid obstacles, reflecting various regime references, (Gutiérrez & Muñoz-Cadena, 2025). Performance metrics include corruption, detection price, compliance cost, and machine adaptability. The experimental design also includes ethical protection measures to deal with the risks of algorithmic bias and to ensure alignment with OECD guidelines on automated systems (Kaye et al., 2023; Odilla, 2024), AI integrity principles (Nichols, 2025). The simulation is administered several times to ensure adequate statistical power, and the outcomes are analyzed using comparative visualization and sensitivity evaluation to identify the key drivers of corruption and device efficacy.

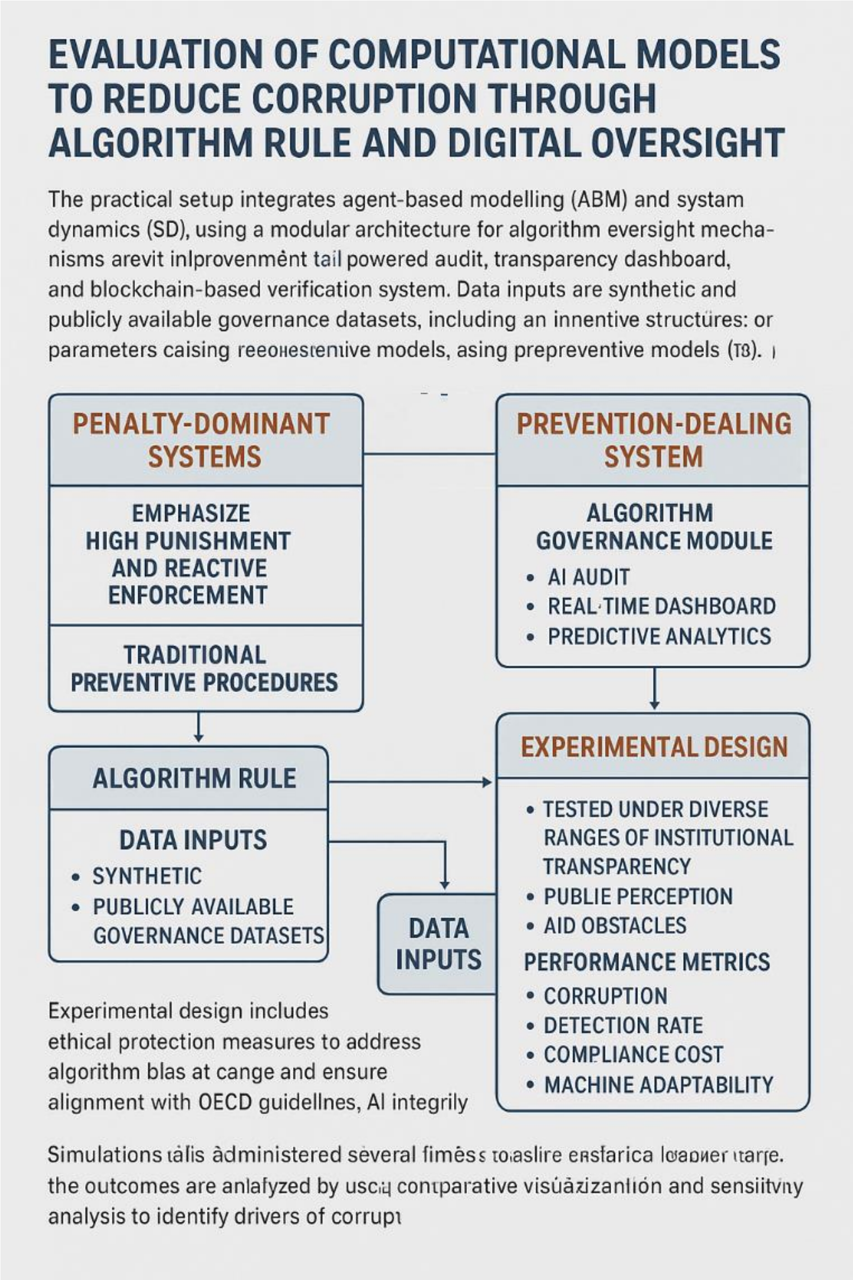

Figure 9. Experimental Setup

7. Results and Analysis

Simulation experiments provided comparative insight into punishment-headed and preventive governance systems. The findings suggest that rule-based frameworks and embedded digital oversight will substantially increase systemic performance and substantially reduce the incidence of corruption.

7. 1. Decreased Corruption Incident

The penalty-dominant approach, which relies on prevention through high punishment and reactive enforcement, showed short-term period cuts in corruption but did not keep compliance beneath distinct institutional conditions (Kasa et al., 2023). In contrast, prevention-rich models incorporating audit mechanisms, transparency dashboards, and blockchain verification demonstrated a continuous decline in corruption rates (Ayobami et al., 2023; Ibrahimy et al., 2024). Agent-based results revealed a 30–45% reduction in corruption incidents under preventive rule in comparison to punishment-based systems, supporting existing conclusions on the effectiveness of inspection and transparency mechanisms.

7. 2. Effectiveness in Detecting Remote Locations

The algorithmic module improved both the accuracy and timeliness of detection. The AI-operated audit diagnosed discrepancies in procurement and administrative selections at significantly higher rates than conventional tracking (Ayobami et al., 2023; Rysin & Sukh, 2024). The transparency dashboard featured real-time tracking, leaving the "C programming language impact" inherent in punishment-based structures. These outcomes verify the emphasis on OECD's active AI-operated integrity mechanisms (Nichols, 2025) and align with practical implementations such as Ukraine's Prizoro platform (Rysin & Sukh, 2024).

7. 3. Compliance Fee and Performance

While punishment-based tactics imposed a good enough compliance fee on establishments and citizens, preventive structures demonstrated greater efficiency. The simulation output showed that the virtual oversight reduced the executive burden through detecting and reducing unnecessary punitive capabilities (Gaurav et al., 2025; Srinivasan & Patapati, 2025). Blockchain-enabled traceability further reinforced verification tactics (Ibrahimy et al., 2024). These findings are consistent with prior studies that highlight the efficiency benefits of rules-based integrity frameworks in public procurement (Ayobami et al., 2023).

7. 4. Public Trust and Legitimacy

Public trust plays a vital role in the adoption of governance systems. The findings indicate that transparency dashboards and audit trails significantly multiplied the belief in oversight mechanisms, thereby reducing the disintegration of rule-based systems related to corruption perceptions (Çani et al.; Castelo, 2024). However, issues regarding potential prejudice in AI-operated systems persist (Odilla, 2024), highlighting the need for impartial system design and proper human oversight.

7. 5. Moral and Institutional Views

Simulation results additionally highlighted the possible risks of over-reliance on automated systems, especially in low-answer or politically sensitive contexts (Aarvik, 2019; Gutiérrez & Muñoz-Cadena, 2025; Kaye et al., 2023). While preventive systems advanced beyond punitive models, their effectiveness depended on ethical protection measures, algorithmic transparency and institutional alignment. These findings enhance the broad theoretical understanding of the moral basis and enduring transformations of anticorruption and permanent change.

Overall, the findings support the view that a shift from fines-based prevention to prevention-based regulation within a rules-driven regime leads to improved integrity, performance and public trust. However, the integration of equity, transparency, and ethical inspection is vital to avoid exacerbating systemic inequalities or algorithmic bias.

Figure 10. Results and Analysis

8. Discussion

The findings of this study underscore the transformative potential of rule-based systems and digital inspection to reduce corruption through preventive measures rather than punitive mechanisms. Simulation results indicate that AI-powered audits, transparency dashboards, and blockchain-based verification structures not only reduce the prevalence of low corruption but also enhance the performance of detection, lower compliance costs, and promote public trust. These outcomes align with rising paradigms within the AI regime, including rule-based governance and democracy-in-silico, which emphasize the position of automated systems in strengthening institutional integrity.

However, adopting such structures is not without challenges. Common threats include algorithmic bias, overlapping on automation, and interpretative obstacles, which highlight the need for robust moral protection measures and transparent design principles (Kaye et al., 2023; Odilla, 2024). An examination of public perception shows that trust in monitoring algorithms is closely linked to fairness and accountability, both of which ensure technical effects. In addition, the scalability of these models in various institutional contexts remains an open question, particularly in resource-constrained or politically volatile environments. Lessons from early implementation in Latin America and Ukraine indicate that success often relies on complementary institutional reforms, potential construction and adaptive governance structures (Gutiérrez & Muñoz-Cadena, 2025; Rysin & Sukh, 2024)[9]. This research contributes to the field of AI morality and the large discourse on governance by providing a computational framework that integrates technological innovation with authentic ideas. It also offers practical insights for policymakers seeking to transition from reactive punishment to proactive, digitally enabled oversight systems. Future research should focus on evaluating real-world pilot implementations, human-AI collaborative models, and emerging technology, such as decentralized self-sustaining agencies (DAOs) and real-time predictions, to similarly enhance the flexibility and equity of the anti-corruption gadget.

9. Conclusion

The study presents a computational model that integrates algorithmic regime and digital oversight to address corruption by shifting from reactive, punishment-based anti-corruption strategies to prevention-oriented structures. Using agent-based and machine dynamics simulations, the model indicates that algorithmic oversight, along with AI-powered auditing, transparency dashboards, and blockchain verification, reduces corruption incidence, enhances detection capabilities, lowers compliance costs, and improves public trust. These conclusions underline the capacity of virtual devices to reform governance structures by embedding integrity, transparency, and performance into institutional strategies. However, adopting such systems is not without challenges. Ethical risks, including algorithmic bias, overreliance on automation, and lecturer boundaries, need to be managed cautiously through human-centred design, third-party audits, and continuous human oversight. The success of the algorithmic regime depends on extensive institutional reforms, public values and its alignment with well-known outlines such as constitutional AI and moral public management.

10. Future Work

Must focus on several key areas:

- Blockchain Integration: Exploring decentralized account technology for irreversible audit trails and extended transparency in public transactions.
- Real-Time Monitoring: Developing a dynamic, adaptive oversight gadget capable of responding to the risks of emerging corruption in real time.
- Human-AI Cooperation: Designing hybrid governance models that leverage the strengths of both rule-based systems and human decision-making to ensure accountability and appropriate contextual sensitivity.
- Cross-Cultural Verification: Testing models across diverse institutional and cultural settings to assess their scalability and impact.
- Ethical AI Regime: Conducting research at the design level that encompasses risk mitigation, interpretable AI, and equitable, legitimate virtual oversight.

By bridging computational innovation with ethical and institutional hardness, this study provides a pathway to more flexible, transparent, and effective governance structures capable of reducing corruption through prevention rather than punishment.

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Data Availability

No data was used for the research described in the article.

References

Aarvik, P. (2019). Artificial Intelligence–a promising anti-corruption tool in development settings. U4 Anti-Corruption Resource Centre.

Abdulwasaa, M. A., Kawale, S. V., Abdo, M. S., Albalwi, M. D., Shah, K., Abdalla, B., & Abdeljawad, T. (2024). Statistical and computational analysis for corruption and poverty model using Caputo-type fractional differential equations. Heliyon, 10(3).

Ayobami, A. T., Mike-Olisa, U., Ogeawuchi, J. C., Abayomi, A. A., & Agboola, O. A. (2023). Algorithmic Integrity: A Predictive Framework for Combating Corruption in Public Procurement through AI and Data Analytics. Journal of Frontiers in Multidisciplinary Research, 4(2), 130-141.

Çani, E. M., Çuka, M., & Mazelliu, A. Transparency and Accountability in AI Systems: A Realistic Approach in Albania. In Artificial Intelligence in Legal Systems (pp. 50-62). Chapman and Hall/CRC.

Castelo, N. (2024). Perceived corruption reduces algorithm aversion. Journal of Consumer Psychology, 34(2), 326-333.

Ceva, E., & Jiménez, M. C. (2022). Automating anticorruption? Ethics and Information Technology, 24(4), 48.

Elnawawy M, S., Okasha, A. E., & Hosny, H. A. (2022). Agent-based models of administrative corruption: An overview. International Journal of Modelling and Simulation, 42(2), 350-358.

Gaurav, S., Heikkonen, J., & Chaudhary, J. (2025). Governance-as-a-service: A multi-agent framework for ai system compliance and policy enforcement. arXiv preprint arXiv:2508.18765.

Gutiérrez, J. D., & Muñoz-Cadena, S. (2025). Artificial intelligence in Latin America's public policy cycles. In Handbook of Public Policy in Latin America (pp. 405-422). Edward Elgar Publishing.

Ibrahimy, M. M., Norta, A., & Normak, P. (2024). Blockchain-based governance models supporting corruption-transparency: A systematic literature review. Blockchain: Research and Applications, 5(2), 100186.

Kasa, R., Réthi, G., Hauber, G., & Szegedi, K. (2023). Simulation of corruption decisions—an agent-based approach. Sustainability, 15(3), 2561.

Kaye, K., Dixon, P., & Gellman, R. (2023). Risky Analysis: Assessing and Improving AI Governance Tools. An International Review of AI Governance Tools and Suggestions for Pathways Forward.

Nichols, P. M. (2025). Does compliance with the global anticorruption regime require the use of artificial intelligence? American Business Law Journal, 62(3), 145-164.

Odilla, F. (2024). Unfairness in AI Anti-Corruption Tools: Main Drivers and Consequences: Unfairness in AI Anti-Corruption Tools. Minds and Machines, 34(3), 28.

Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C. J., Morris, M. R., Willer, R., Liang, P., & Bernstein, M. S. (2025). Simulating human behavior with ai agents. In.

Petheram, A., Pasquarelli, W., & Stirling, R. (2019). The next generation of anti-corruption tools: Big data, open data & artificial intelligence. Available at researchreport2019_thenextgenerationofanti-corruptiontools_bigdataopendataartificialintelligence. pdf (europa. eu)(Accessed 28 February 2024).

Raji, I. D., Xu, P., Honigsberg, C., & Ho, D. (2022). Outsider oversight: Designing a third party audit ecosystem for ai governance. Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society,

Rysin, M., & Sukh, Y. (2024). Digital solutions as an effective approach to combat corruption in public procurement. Democratic Governance, 2(17), 18-29.

Schmitz, C., Rystrøm, J., & Batzner, J. (2025). Oversight structures for agentic AI in public-sector organizations. Proceedings of the 1st Workshop for Research on Agent Language Models (REALM 2025),

Srinivasan, T., & Patapati, S. (2025). Democracy-in-Silico: Institutional Design as Alignment in AI-Governed Polities. arXiv preprint arXiv:2508.19562.