IIAR JOURNALS

Innovating Knowledge. Advancing Research.

IIAR Journal of Business, Management, and Entrepreneurship

Vol 1, No 1, 2026

Responsible AI Governance in the Firm: A Conceptual Framework Linking Board Oversight, Accountability Structures and Organisational Legitimacy

NWOGBOJI Emmanuel Usulor[1], OGAH Sunday Ogbo[2]

[1,2] Veristzon Global Research Lab

ABSTRACT

Artificial intelligence (AI) is now used in pricing, hiring, lending, customer service and strategy, and the governance of these systems has become a board-level responsibility. Yet most governance scholarship treats AI as a technology risk to be delegated to management rather than as an accountability question for directors, and the growing literature on responsible AI pays little attention to the board. At the same time, firms increasingly publish AI principles whose credibility is contested, raising the question of when AI governance is substantive and when it is merely symbolic. This conceptual paper develops a framework for responsible AI governance within the firm. Methodologically, it uses a theory-synthesis design based on an integrative review of 43 peer-reviewed journal articles published between 2023 and 2026. Each article was verified against its bibliographic record and screened for retractions and corrections, and the articles were coded thematically by governance problem, board function and governance mode. The synthesis integrates corporate governance theory, stakeholder and legitimacy theory, and research on AI governance, accountability and organisational AI ethics. It identifies four distinctive governance problems created by AI: opacity, scale, drift and diffuse responsibility. It maps each problem onto the board functions of strategy oversight, risk oversight and accountability. It distinguishes symbolic from substantive AI governance and derives six propositions on the roles of board AI expertise, the verifiability of institutional pressure, integration into existing risk and audit structures, and institutional voids. The paper argues that substantive AI governance is a source of durable organisational legitimacy with investors, regulators and customers, rather than a compliance cost, whereas symbolic governance offers fragile legitimacy that can collapse after an AI incident. Situating the framework in emerging-economy governance conditions, it recommends that Nigerian and African boards treat AI as a standing strategic and risk matter and build or access AI expertise. Boards should also assign AI oversight to existing committees, designate accountable executives and report against measurable indicators. Regulators and professional bodies should pair AI disclosure with verification and director education.

Keywords: AI governance; corporate governance; board oversight; accountability; responsible AI

1. INTRODUCTION

Artificial intelligence (AI) is no longer confined to experimental units of the firm. Algorithms now set prices, screen job applicants, score borrowers, answer customers and inform strategic choices, and generative systems have widened the range of tasks that can be delegated to machines (Taeihagh, 2025). As deployment has spread, so has evidence that AI creates harms that are difficult to see, quick to multiply and hard to attribute to any one person. These features turn AI from a narrow technical matter into a question of corporate accountability: when an automated decision harms a customer, misleads an investor or breaches a regulation, someone in the firm must be answerable for it (Novelli et al., 2024).

Corporate governance scholarship has been slow to absorb this shift. A large literature on responsible AI has produced principles, toolkits and audit procedures (Lu et al., 2024; Laine et al., 2024), and systematic reviews now map organisational AI governance in detail (Birkstedt et al., 2023; Batool et al., 2025; Papagiannidis et al., 2025). Yet most of this work locates governance in management, engineering teams or ethics committees, and treats the board, at most, as a distant recipient of reports. Birkstedt et al. (2023) note that implementation and senior-level commitment remain among the least understood aspects of AI governance. Legal and governance scholars have only recently begun to ask what AI means for directors' duties and board processes (Brand, 2024; Agnese et al., 2025; Ben-Shahar et al., 2026).

A second problem is credibility. Firms have published AI principles in large numbers, but critics argue that many commitments are vague, unenforced and convenient for corporate interests (Munn, 2023). The term "ethicswashing" now describes communication that makes AI practices appear ethical without substantive backing (Schultz et al., 2025), and philosophers have asked how outsiders can tell a serious ethics office from a symbolic one (Biddle et al., 2024). Recent analyses of corporate AI principles show that they function partly as legitimation devices aimed at investors, customers and regulators (Huhta et al., 2026; Zhang et al., 2026). What the literature lacks is an account that connects these legitimacy dynamics to the formal machinery of corporate governance: the board's oversight of strategy and risk and its duty to render account.

This gap matters especially for emerging economies. In Nigeria and much of Sub-Saharan Africa, AI policy frameworks are recent and uneven (Ayana et al., 2024; Plantinga et al., 2024), regulatory enforcement is often weak, and board behaviour is shaped by informal institutions as much as by formal codes (Nakpodia et al., 2025a). Firms in such settings increasingly import AI systems built elsewhere, while their boards face investors, partners and regulators who expect credible assurance that those systems are under control.

This paper therefore addresses the question: *how should responsible AI governance be conceptualised as a board-level accountability function, and under what conditions do firms adopt substantive rather than symbolic AI governance?* Its objectives are (i) to specify the governance problems that distinguish AI from other technologies; (ii) to map those problems onto the core board functions of strategy oversight, risk oversight and accountability; (iii) to distinguish symbolic from substantive AI governance and derive testable propositions about their antecedents and consequences; and (iv) to draw out implications for boards, regulators and professional bodies in emerging economies, with particular reference to Nigeria.

2. LITERATURE REVIEW

2.1 Conceptual Review

2.1.1 AI governance and responsible AI

Organisational AI governance has been defined as the rules, practices and processes that align an organisation's use of AI with its strategy, values, legal obligations, ethical principles and stakeholder expectations (Birkstedt et al., 2023). Responsible AI governance extends this definition by specifying the structural, relational and procedural practices through which responsible-AI principles are translated into the design, operation, monitoring and evaluation of AI systems (Papagiannidis et al., 2025). Across reviews, recurring governance dimensions include accountability, transparency, explainability,

© IIARJournal. All rights reserved.

fairness, privacy, safety and security (Camilleri, 2024; Díaz-Rodríguez et al., 2023). Governance artefacts operate at team, organisational, industry, national and international levels (Batool et al., 2025; Lu et al., 2024). This paper focuses on the organisational level and, within it, on the board.

2.1.2 Board oversight

Boards are conventionally charged with overseeing strategy, overseeing risk and holding management to account. Applied to AI, these functions imply that directors must understand how AI affects the firm's business model, must be satisfied that AI risks are identified and controlled, and must ensure that the firm can answer for AI-driven outcomes. Agnese et al. (2025) argue that AI alters the board's capabilities, roles and functions and requires directors with technical understanding, supported by a specialised AI committee. Brand (2024) contends that directors' duties are an under-used instrument of AI regulation precisely because boards shape how AI is developed and adopted.

2.1.3 Accountability

Accountability is understood here as answerability: an agent's obligation to explain and justify conduct to a forum that can pass judgement and impose consequences. Novelli et al. (2024) decompose AI accountability into seven features (context, range, agent, forum, standards, process and implications) and four goals (compliance, reporting, oversight and enforcement). This decomposition is useful for corporate governance because it separates *who* is answerable (the agent) from *to whom* (the forum), a distinction that AI systems tend to blur.

2.1.4 Symbolic and substantive governance

Symbolic governance refers to structures and statements that signal conformity with expectations without materially altering practice; substantive governance changes resource allocation, decision rights and behaviour. The AI ethics literature describes the symbolic form through the vocabulary of ethicswashing (Schultz & Seele, 2023; Schultz et al., 2025) and "toothless" principles (Munn, 2023). Biddle et al. (2024) provide a test grounded in organisational power: an ethics function is substantive to the extent that it has the authority and resources to change outcomes.

2.1.5 Organisational legitimacy

Organisational legitimacy is the perception among relevant audiences that a firm's actions are proper and appropriate. Firms seek it through legitimation, which may rely on symbolic cues, such as statements of values, or substantive cues, such as verifiable practices (Zhang et al., 2026). For AI, the relevant audiences include investors assessing material risk (Minkkinen et al., 2024), regulators enforcing new risk-based regimes (Schuett, 2024), and customers and employees whose trust depends on whether decisions can be explained (Grimmelikhuijsen, 2023).

2.2 Theoretical Review

2.2.1 Agency theory and directors' duties

Agency theory frames the board as a monitor of managerial agents on behalf of shareholders. AI adds a second layer of delegation: managers delegate decisions to systems whose reasoning they may not fully understand. The board therefore monitors an agent who is itself monitoring an opaque artefact. Brand (2024) argues that directors' duties suit this problem because they rely on flexible standards and after-the-event assessment of complex decisions. Agency theory also warns that information asymmetry between management and the board will be acute for AI, where technical knowledge sits deep in the organisation (Agnese et al., 2025).

2.2.2 Stakeholder theory

Stakeholder theory holds that the firm is accountable to all groups that affect or are affected by its activities. AI widens the set of affected parties, because one model can touch millions of customers, applicants or citizens simultaneously. Schultz and Seele (2023) argue that business ethics experience in stakeholder management, standardised reporting and corporate governance can be transferred to organisational AI ethics. Camilleri (2024) similarly concludes that all actors who develop or maintain AI systems carry social responsibilities towards stakeholders.

2.2.3 Legitimacy and institutional theory

Institutional theory predicts that organisations adopt structures that are valued in their environment, sometimes "decoupling" those structures from core practice when conformity is costly. Corporate AI principles fit this pattern: Huhta et al. (2026) show

© IIARJournal. All rights reserved.

that the AI principles of large firms work as institutional messages that legitimise firms as "AI enablers". Zhang et al. (2026), applying legitimacy theory to corporate ethics statements, find both symbolic and substantive legitimation cues. Together these theories suggest that AI governance can be adopted for its signalling value, for its control value, or both, and that which motive dominates depends on whether audiences can verify the difference.

2.3 Empirical Review

Empirical work on organisational AI governance is growing but remains dominated by case studies and document analyses. Papagiannidis et al. (2023), studying three Norwegian energy firms, identify governance practices that help organisations overcome barriers to realising value from AI. Mökander and Floridi (2023), in a year-long study of an ethics-based AI audit at AstraZeneca, find that the main difficulties were classic governance problems: harmonising standards across decentralised units, defining audit scope, managing internal communication and measuring outcomes. Sanderson et al. (2023), interviewing AI designers and developers, document a persistent gap between high-level principles and day-to-day practice.

Studies of disclosure reveal how firms present AI governance externally. Bonsón et al. (2023) analysed 962 annual and sustainability reports of 337 listed Western European firms and found that disclosure on automated decision-making was at an early stage, led by the financial sector. In an emerging-market setting, Elsayed and Ahmed (2025) found that larger boards and more gender-diverse boards were associated with greater AI disclosure in Egyptian banks. Galani et al. (2026) report that AI disclosure among large global corporations accelerated after 2023 but was not associated with short-term financial or employment outcomes, suggesting that disclosure reflects communication more than verified adoption. Investors themselves have limited knowledge of AI and rarely integrate it systematically into ESG analysis (Minkkinen et al., 2024), although protocols for evaluating and disclosing AI's ESG implications have been proposed (Sætra, 2023).

Evidence on boards is thinner still. Ahdadou et al. (2024, 2026) analyse stages of AI integration in boardroom decision-making and argue that augmented, human-in-the-loop configurations are the most viable. Ben-Shahar et al. (2026) show theoretically that AI in the boardroom can degrade information heterogeneity and produce "algorithmic groupthink". In Nigeria, qualitative studies show that board processes are shaped by informal logics such as ethnicity (Nakpodia et al., 2025a), that subsidiaries of multinational firms face low-visibility institutional challenges (Nakpodia et al., 2025b), and that shareholder activism operates differently in weak institutional settings (Ashiru et al., 2023). At the policy level, African AI strategies are often drafted without assessing local implementability (Plantinga et al., 2024), data governance needs context-specific design (Effoduh et al., 2024), and progress towards locally grounded AI governance is uneven across Sub-Saharan countries (Ayana et al., 2024).

2.4 Research Gap

Three gaps follow. First, AI governance research rarely specifies why AI requires board attention beyond that given to other technologies; its distinctive properties are discussed separately rather than mapped onto board functions. Second, the symbolic–substantive distinction is well developed in AI ethics but is not connected to governance structures or to the conditions that favour one mode over the other. Third, emerging-economy governance conditions, where enforcement is weak and imported AI is common, are almost absent from the board-level AI literature. This paper addresses all three.

3. MATERIALS AND METHODS

3.1 Research Design

The study is conceptual and non-empirical. It follows a theory-synthesis design: it integrates findings and constructs from separate literatures (corporate governance, AI governance and accountability, organisational AI ethics, and legitimacy and institutional theory) into a framework with explicit propositions that can later be tested empirically. No primary data were collected and no human participants were involved.

3.2 Population and Sample of Literature

The population consisted of peer-reviewed journal articles on organisational AI governance, AI accountability, board oversight of AI, AI disclosure and legitimacy, distinctive AI risks, and AI and

© IIARJournal. All rights reserved.

corporate governance in African and emerging-economy contexts. Discovery used web-indexed scholarly search across publisher platforms and institutional repositories, with targeted searches on each theme and backward checks of key reviews. About 75 candidate records were screened at full-record level, and 43 articles were retained for the final synthesis.

3.3 Inclusion and Exclusion Criteria

Articles were included if they (a) were published in a peer-reviewed journal; (b) had a publication year between 2023 and 2026, taking the issue year or, for advance online articles, the online year; and (c) addressed at least one of the themes above. Preprints, conference proceedings, books, working papers, practitioner reports and law-firm commentaries were excluded, as were items whose bibliographic record could not be confirmed. Where an article had a journal version and a preprint, only the journal version was cited.

3.4 Data Collection and Verification

Each retained record was checked against a primary bibliographic source, namely the publisher's page, Crossref or OpenAlex metadata, PubMed, or an institutional repository record. The check covered authors, title, journal, year, volume, issue, pages or article number, and DOI. Records were also screened for retractions, corrections and expressions of concern. No retracted item was retained. Two retained articles carry minor published corrections (Mökander & Floridi, 2023; Ahdadou et al., 2024), neither affecting the claims for which they are cited. Two candidates whose volume or pages could not be confirmed were dropped.

3.5 Data Analysis and Bases of Judgement

Articles were coded thematically in three passes. The first pass coded the governance problems attributed to AI. The second coded governance responses by the board function they serve. The third coded evidence on the antecedents and consequences of symbolic or substantive governance. Propositions were derived where at least two independent sources supported a relationship, or where a relationship followed logically from the integrated theories. Judgements were based on theoretical coherence, consistency with the reviewed evidence, and testability.

3.6 Reliability and Validity

Reliability was supported by a transparent protocol: fixed inclusion rules, record-level verification and explicit coding categories, all reported here. Construct validity was addressed by anchoring each construct in definitions from the reviewed literature (Table 1). Because the framework is conceptual, its external validity remains to be established. The propositions are therefore framed as testable relationships, with indicative measures for each construct.

3.7 Constructs and Indicative Measurement

Table 1 defines the framework's constructs and proposes indicative measures for future empirical work.

Table 1. Core constructs, definitions and indicative measures

Construct	Definition	Indicative measures
AI governance exposure	Degree of opacity, scale, drift and diffuse responsibility in the firm's AI portfolio	Inventory of AI systems by risk tier; share of decisions automated; vendor-supplied model share
Board AI oversight	Allocation of AI matters to the board and its committees	AI in committee charters; frequency of AI agenda items; AI reports to board
Board AI expertise	Directors' knowledge of AI and data	Directors with AI/data background; AI training hours; access to independent advisers
Substantive AI governance	Coupling of AI commitments to decision rights, resources, monitoring and consequences	Veto/escalation rights; dedicated budget; independent audit; KPIs; incident reporting
Symbolic AI governance	Visible AI commitments decoupled from practice	Principles without procedures; boilerplate disclosure; ethics bodies without authority
Verifiability of pressure	Extent to which audiences can check AI governance in operation	Regulatory audit or conformity duties; investor AI metrics; analyst/media scrutiny
Organisational legitimacy	Audience judgements that the firm's AI use is proper and appropriate	Investor ESG ratings; regulatory sanctions; customer trust and complaint data

Note. Measures are indicative and intended for future empirical testing. Source: Author's synthesis of the reviewed literature.

3.8 AI Usage Disclosure

A generative AI assistant was used to support literature retrieval, bibliographic verification and drafting. All sources were checked against primary bibliographic records as described above. The author(s) reviewed and edited the full text and accept responsibility for its content.

© IIARJournal. All rights reserved.

4. RESULTS AND DISCUSSION

4.1 RESULTS

4.1.1 Four distinctive governance problems of AI

The synthesis identifies four properties that together distinguish AI from earlier enterprise technologies and justify board-level attention.

Opacity. Many AI models, particularly machine-learning and generative systems, do not produce reasons in a form that humans can readily inspect (Taeihagh, 2025). Opacity weakens the board's capacity to evaluate management's claims about AI and the firm's capacity to explain decisions to affected people. Experimental evidence shows that explainability, more than mere access to an algorithm, shapes whether automated decisions are perceived as trustworthy (Grimmelikhuijsen, 2023).

Scale. A single model can make millions of decisions, so an error or bias is replicated instantly across customers, applicants or markets. Risk-based regulation ties obligations to the context and consequences of use (Schuett, 2024), although critics question whether a regulatory judgement of acceptable risk makes a system trustworthy (Laux et al., 2024). Scale converts what would be isolated human errors into systemic exposure that can become material to the firm.

Drift. AI performance is not fixed at deployment. As the data a model encounters diverge from its training data, its performance can deteriorate without visible failure (Sahiner et al., 2023). Drift means that approval at launch provides no lasting assurance. Governance must be continuous, with monitoring, re-validation and escalation triggers across the system's life.

Diffuse responsibility. AI systems are produced and operated by many hands: data suppliers, vendors, developers, business owners and users. The philosophical debate on "responsibility gaps" shows that it can be unclear who is answerable for an AI-driven outcome (Da Silva, 2024; Oimann & Tollon, 2025), and that such gaps often reduce to problems of collective agency (Llorca Albareda, 2025). Inside firms, responsibility for AI risk is frequently unassigned or fragmented (Schuett, 2025).

4.1.2 Mapping the problems onto board functions

Table 2 maps each problem onto the three board functions and identifies the governance mechanisms through which the board can respond. Two patterns stand out. First, every problem has implications for all three functions, which is why AI governance cannot be fully delegated to a technical committee. Second, the problems interact: opacity makes drift harder to detect, scale magnifies the consequences of undetected drift, and diffuse responsibility means that no one may be watching.

To illustrate, consider a hypothetical bank that deploys a vendor-supplied credit-scoring model across its retail portfolio. Opacity means that loan officers cannot explain declines to applicants. Scale means that any bias against a region or group is applied to every application. Drift means that a model calibrated before an economic shock may quietly misprice risk afterwards. Diffuse responsibility means that the vendor, the data team, the credit function and compliance can each assume that another is watching. A board that receives only a statement that the bank "uses AI responsibly" has no basis for oversight. A board that has approved the deployment against a risk appetite, receives drift and fairness indicators, and can name the executive accountable for the model is exercising all three functions.

Table 2. Mapping AI governance problems onto board functions

Problem	Strategy oversight	Risk oversight	Accountability
Opacity	Set explainability thresholds by use case; avoid opaque models where reasons are owed	Require independent model validation and documentation	Disclose AI use; provide explanation and contestation routes
Scale	Approve high-impact deployments; tier the AI portfolio by risk	Conduct impact assessments; set AI risk appetite and exposure limits	Provide redress at scale; report aggregate outcomes
Drift	Fund lifecycle monitoring, not only deployment	Monitor continuously; set key risk indicators and re-validation triggers; include in internal audit	Report periodically to the board; disclose incidents
Diffuse responsibility	Name an accountable executive for each high-impact system	Assign roles across the three lines of defence	Define the AI remit in committee charters; link to performance evaluation

© IIARJournal. All rights reserved.

Source: Author's synthesis drawing on Novelli et al. (2024), Schuett (2024, 2025), Agnese et al. (2025) and Papagiannidis et al. (2025).

4.1.3 Symbolic versus substantive AI governance

Table 3 operationalises the distinction between symbolic and substantive AI governance across six dimensions. The defining feature of substantive governance is coupling: commitments are connected to decision rights, resources, monitoring and consequences. Following Biddle et al. (2024), the critical test is power, meaning whether the responsible function can delay, change or stop a deployment.

Table 3. Symbolic versus substantive AI governance

Dimension	Symbolic (decoupled)	Substantive (coupled)
Principles	Published values with no operating procedures	Principles translated into policies, standards and controls
Structure	Advisory ethics body outside decision lines	Board committee charter; accountable executive; three lines of defence
Power	No authority to delay or stop deployments	Escalation and veto rights over high-risk uses
Resources	No dedicated budget or staff	Funded monitoring, validation and training
Assurance	Self-declared compliance	Independent internal or external audit
Disclosure	Narrative and boilerplate statements	Inventories, metrics, incidents and remediation

Source: Author's synthesis drawing on Munn (2023), Biddle et al. (2024), Schultz et al. (2025) and Zhang et al. (2026).

4.1.4 Propositions

The framework yields six propositions, summarised in Table 4 and depicted in Figure 1.

Proposition 1 (problem exposure). The greater the opacity, scale, drift and diffusion of responsibility in a firm's AI portfolio, the greater the accountability gap created when AI governance is delegated wholly to management. Firms with high exposure and no board-level oversight will experience more severe AI-related incidents and legitimacy losses. This follows from the problems described in Section 4.1.1. Each erodes the information on which delegated monitoring depends, which is the agency logic of Section 2.2.1.

Proposition 2 (board AI expertise). Board AI expertise is positively associated with substantive rather than symbolic AI governance. Directors with technical understanding can interrogate management, reduce information asymmetry and recognise empty commitments (Agnese et al., 2025). Board composition is already associated with the extent of AI disclosure (Elsayed & Ahmed, 2025). Expertise should also guard against the epistemic capture that unreflective use of AI in the boardroom can produce (Ben-Shahar et al., 2026).

Proposition 3 (verifiability of pressure). Institutional pressures that reward the visibility of AI governance, but cannot verify its operation, increase the likelihood of symbolic governance. Pressures that come with verification increase the likelihood of substantive governance. Examples of verifying pressures include regulatory conformity assessment, independent audit and investor engagement using defined metrics. Where audiences see only statements, firms gain legitimacy at low cost by publishing principles (Huhta et al., 2026; Galani et al., 2026). Where audits and risk-management obligations apply, decoupling becomes costly (Schuett, 2024; Mökander et al., 2024).

Proposition 4 (integration into existing structures). AI governance that is integrated into existing enterprise risk management, internal audit and board-committee structures is more likely to be substantive than governance housed in a stand-alone ethics body. Integration assigns roles across management, risk and compliance, and internal audit (Schuett, 2025). It also links AI oversight to established reporting lines to the board, avoiding the isolation that weakens symbolic ethics offices (Biddle et al., 2024).

Proposition 5 (legitimacy consequences). Substantive AI governance generates more durable legitimacy with investors, regulators and customers than symbolic governance. Symbolic governance may yield short-term legitimacy, but it carries a legitimacy discount when an AI incident exposes the gap between statement and practice. The ethicswashing literature describes this reversal (Schultz et al., 2025), and legitimacy research shows that audiences distinguish symbolic from substantive cues (Zhang et al., 2026). Investors' demand for standardised AI metrics is likely to sharpen the distinction over time (Minkkinen et al., 2024; Sætra, 2023).

Proposition 6 (institutional voids). In emerging economies with weak external enforcement, board-level AI governance partly substitutes for absent regulatory assurance. In such settings, substantive governance yields a larger legitimacy premium,

© IIARJournal. All rights reserved.

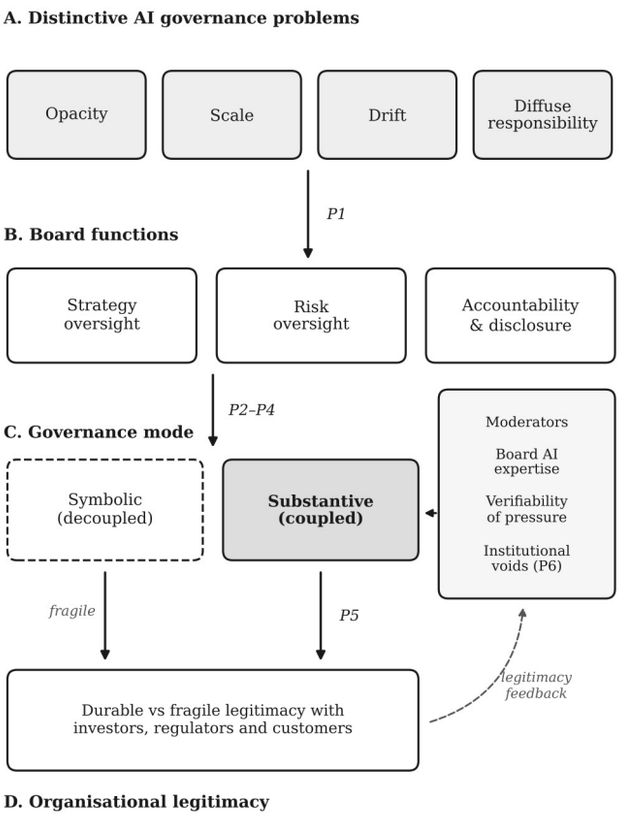
particularly with foreign investors, partners and regulators. Where AI policy frameworks are immature and capacity is limited (Ayana et al., 2024; Plantinga et al., 2024), credible signals of control must come from the firm. Corporate governance mechanisms such as stakeholder engagement and accountability drivers have been shown to address institutional challenges in Nigeria (Nakpodia et al., 2025b).

Table 4. Summary of propositions

No.	Proposition	Theoretical basis
P1	High AI governance exposure without board oversight widens the accountability gap and raises incident and legitimacy risk	Agency theory
P2	Board AI expertise is positively associated with substantive governance	Agency theory; directors' duties
P3	Visibility-only pressures favour symbolic governance; verifying pressures favour substantive governance	Institutional theory
P4	Integration into enterprise risk management, audit and committee structures favours substantive governance	Agency theory; risk governance
P5	Substantive governance yields durable legitimacy; symbolic governance carries a legitimacy discount when exposed	Legitimacy and stakeholder theory
P6	In institutional voids, substantive board-level AI governance substitutes for weak enforcement and earns a legitimacy premium	Institutional and legitimacy theory

Source: Author's conceptualisation.

Figure 1. Conceptual framework of responsible AI governance in the firm



Note. P = proposition. Dashed box = symbolic governance.
Source: Author's conceptualisation.

4.2 DISCUSSION

4.2.1 Relation to prior studies and theory

The framework extends prior reviews of AI governance in three ways. Birkstedt et al. (2023) and Papagiannidis et al. (2025) catalogue governance themes and practices but do not specify why particular practices are needed at board level. The mapping of opacity, scale, drift and diffuse responsibility supplies that rationale and links each problem to strategy oversight, risk oversight and accountability. It thereby answers calls for research on AI oversight units and senior-management commitment (Birkstedt et al., 2023). The framework is consistent with Agnese et al. (2025) and Brand (2024), who argue that AI changes what boards must know and do, but it moves beyond them by connecting board design to the credibility of AI governance.

The framework also gives the critical AI ethics literature an institutional explanation. Munn (2023)

© IIARJournal. All rights reserved.

and Schultz et al. (2025) describe ethicswashing as a pervasive problem. Proposition 3 explains its prevalence as a rational response to audiences that cannot verify practice, and Proposition 5 predicts its fragility. This reading is consistent with case evidence that operationalising AI governance is constrained by familiar organisational factors (Mökander & Floridi, 2023). It is also consistent with the persistent gap between principles and practice reported by Sanderson et al. (2023).

Theoretically, the framework combines agency, stakeholder and legitimacy logics rather than choosing among them. Agency theory explains why delegation fails under opacity and drift. Stakeholder theory explains why the forum of accountability must extend beyond shareholders when AI operates at scale. Legitimacy theory explains why firms adopt AI governance even when its control value is low, and why that choice is unstable.

4.2.2 Implications for Nigerian and African boards, regulators and professional bodies

For boards, the framework implies five steps. AI should be a standing board agenda item. At least one director should have AI or data expertise, or the board should have structured access to independent advice. Oversight should sit with an existing committee, typically risk or audit, under an explicit charter. A senior executive should be designated as accountable for each high-impact AI system. And AI should be reported against defined indicators rather than narrative assurances. Because many Nigerian firms deploy AI procured from foreign vendors, contracts should secure the documentation, audit and monitoring rights that oversight requires. Given evidence that informal logics shape Nigerian board dynamics (Nakpodia et al., 2025a), formal charters and documented escalation routes matter especially for preventing AI questions from being suppressed in the boardroom.

For regulators and code-setters, Proposition 3 suggests that disclosure requirements alone may encourage symbolic compliance. Disclosure should be paired with verifiable elements: risk-tiered inventories of AI systems, impact assessments for high-risk uses, and independent assurance. Such elements can draw on governance capabilities that already exist in African financial and electronic-transactions regulation (Plantinga et al., 2024). They should also be aligned with continental data-governance instruments (Effoduh et al., 2024).

For professional bodies, director education is the most direct lever on Proposition 2. Institutes of directors, accountants and auditors can develop AI governance curricula and assurance standards that make substantive governance easier to demonstrate. They can also support context-specific approaches to governance in weak institutional settings, a need highlighted for investor activism (Ashiru et al., 2023).

4.2.3 Limitations and future research

The framework is conceptual and draws on a recent, fast-moving literature that is dominated by studies from Europe, North America and Australia. Its propositions require empirical testing. Promising designs include content analysis of board charters and annual reports, combined with incident data, to test Propositions 1, 3 and 5. Survey and archival studies of director expertise could test Proposition 2. Comparative case studies of Nigerian and other African listed firms could test Proposition 6. Future work should also examine how generative AI used by directors themselves affects board deliberation (Ben-Shahar et al., 2026; Ahdadou et al., 2026).

5. CONCLUSION AND RECOMMENDATION

AI raises governance problems that management alone cannot resolve. Its opacity, scale, drift and diffusion of responsibility weaken the information on which delegated monitoring depends, so the board's functions of strategy oversight, risk oversight and accountability must be adapted. Whether firms meet this challenge substantively or symbolically depends on board expertise, on whether external audiences can verify governance, and on whether AI oversight is built into existing risk and audit structures. Substantive governance is a source of durable legitimacy, whereas symbolic governance offers legitimacy that can collapse after a single incident. In emerging economies with institutional voids, credible board-level AI governance can partly substitute for missing regulatory assurance.

The paper recommends that boards (i) treat AI as a standing strategic and risk matter; (ii) build or access AI expertise; (iii) assign AI oversight to an existing committee with a clear charter; (iv) designate

© IIARJournal. All rights reserved.

accountable executives for high-impact systems; and (v) report against measurable indicators. It further recommends that regulators and corporate governance code-setters, including those in Nigeria, pair AI disclosure with verification, and that professional bodies develop director education and assurance standards for AI.

6. CONTRIBUTION TO EXISTING KNOWLEDGE

The paper makes three contributions. First, it maps four distinctive AI problems (opacity, scale, drift and diffuse responsibility) in a way that explains why AI governance is a board-level responsibility and links each problem to specific board functions and mechanisms. Second, it connects the symbolic–substantive distinction from AI ethics to corporate governance structures and derives six testable propositions about the antecedents and legitimacy consequences of each mode. Third, it extends board-level AI governance theory to emerging-economy conditions and identifies concrete implications for Nigerian and African boards, regulators and professional bodies.

7. REFERENCES

Agnese, P., Arduino, F. R., & Di Prisco, D. (2025). The era of artificial intelligence: What implications for the board of directors? *Corporate Governance: The International Journal of Business in Society*, *25*(2), 272–287. https://doi.org/10.1108/CG-06-2023-0259

Ahdadou, M., Aajly, A., & Tahrouch, M. (2024). Unlocking the potential of augmented intelligence: A discussion on its role in boardroom decision-making. *International Journal of Disclosure and Governance*, *21*(3), 433–446. https://doi.org/10.1057/s41310-023-00207-2

Ahdadou, M., Ahdadou, A., Aajly, A., & Tahrouch, M. (2026). Artificial intelligence in corporate boards: A dual-dimensional framework for integration across autonomy and structural levels. *Data Science and Management*, *9*(2), Article 100174. https://doi.org/10.1016/j.dsm.2025.12.001

Ashiru, F., Adegbite, E., & Nakpodia, F. (2023). Nigerian professional investors' sense-making of the impact of shareholder activism on corporate accountability. *Journal of Accounting and Public Policy*, *42*(4), Article 107114. https://doi.org/10.1016/j.jaccpubpol.2023.107114

Ayana, G., Dese, K., Daba Nemomssa, H., Habtamu, B., Mellado, B., Badu, K., Yamba, E., Faye, S. L., Ondua, M., Nsagha, D., Nkweteyim, D., & Kong, J. D. (2024). Decolonizing global AI governance: Assessment of the state of decolonized AI governance in Sub-Saharan Africa. *Royal Society Open Science*, *11*(8), Article 231994. https://doi.org/10.1098/rsos.231994

Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. *AI and Ethics*, *5*, 3265–3279. https://doi.org/10.1007/s43681-024-00653-w

Ben-Shahar, D., Carmeli, A., Sulganik, E., & Weiss, D. (2026). Artificial intelligence transforms board decision-making processes: An extension of the information-processing model of board decision synergy. *AI & Society*, *41*(5), 4349–4364. https://doi.org/10.1007/s00146-025-02773-1

Biddle, J. B., Nelson, J. P., & Olugbade, O. E. (2024). How can we know if you are serious? Ethics washing, symbolic ethics offices, and the responsible design of AI systems. *Canadian Journal of Philosophy*, *54*(4), 329–345. https://doi.org/10.1017/can.2025.9

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research*, *33*(7), 133–167. https://doi.org/10.1108/INTR-01-2022-0042

Bonsón, E., Bednárová, M., & Perea, D. (2023). Disclosures about algorithmic decision making in the corporate reports of Western European companies. *International Journal of Accounting Information Systems*, *48*, Article 100596. https://doi.org/10.1016/j.accinf.2022.100596

Brand, V. (2024). Directors' duties and AI regulation. *Griffith Law Review*, *33*(4), 324–354. https://doi.org/10.1080/10383441.2024.2422221

Camilleri, M. A. (2024). Artificial intelligence governance: Ethical considerations and implications for social responsibility. *Expert Systems*, *41*(7), Article e13406. https://doi.org/10.1111/exsy.13406

Da Silva, M. (2024). Responsibility gaps. *Philosophy Compass*, *19*(9–10), Article e70002. https://doi.org/10.1111/phc3.70002

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy artificial intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, *99*, Article 101896. https://doi.org/10.1016/j.inffus.2023.101896

Effoduh, J. O., Akpudo, U. E., & Kong, J. D. (2024). Toward a trustworthy and inclusive data governance policy for the use of artificial intelligence in Africa. *Data & Policy*, *6*, Article e34. https://doi.org/10.1017/dap.2024.26

Elsayed, D. H., & Ahmed, E. A. (2025). Board of directors' attributes and artificial intelligence disclosure: Evidence from Egyptian banks. *Journal of Financial Reporting and Accounting*. Advance online publication. https://doi.org/10.1108/JFRA-01-2025-0044

Galani, S.-V., Rotaru, G.-C., & Dumitru, A.-M. (2026). AI disclosure dynamics in large global corporations. *Poslovna izvrsnost – Business Excellence*, *20*(1), 43–64. https://doi.org/10.22598/pi-be/2026.1.42656

© IIARJournal. All rights reserved.

Grimmelikhuijsen, S. (2023). Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review*, *83*(2), 241–262. https://doi.org/10.1111/puar.13483

Huhta, M., Sivunen, A., & van Zoonen, W. (2026). Constituting AI-enabling organizations: Legitimation through AI principles. *International Journal of Business Communication*. Advance online publication. https://doi.org/10.1177/23294884261465791

Laine, J., Minkkinen, M., & Mäntymäki, M. (2024). Ethics-based AI auditing: A systematic literature review on conceptualizations of ethical principles and knowledge contributions to stakeholders. *Information & Management*, *61*(5), Article 103969. https://doi.org/10.1016/j.im.2024.103969

Laux, J., Wachter, S., & Mittelstadt, B. (2024). Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk. *Regulation & Governance*, *18*(1), 3–32. https://doi.org/10.1111/rego.12512

Llorca Albareda, J. (2025). Uncovering the gap: Challenging the agential nature of AI responsibility problems. *AI and Ethics*. Advance online publication. https://doi.org/10.1007/s43681-025-00685-w

Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2024). Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. *ACM Computing Surveys*, *56*(7), Article 173. https://doi.org/10.1145/3626234

Minkkinen, M., Niukkanen, A., & Mäntymäki, M. (2024). What about investors? ESG analyses as tools for ethics-based AI auditing. *AI & Society*, *39*(1), 329–343. https://doi.org/10.1007/s00146-022-01415-0

Mökander, J., & Floridi, L. (2023). Operationalising AI governance through ethics-based auditing: An industry case study. *AI and Ethics*, *3*(2), 451–468. https://doi.org/10.1007/s43681-022-00171-7

Mökander, J., Schuett, J., Kirk, H. R., & Floridi, L. (2024). Auditing large language models: A three-layered approach. *AI and Ethics*, *4*(4), 1085–1115. https://doi.org/10.1007/s43681-023-00289-2

Munn, L. (2023). The uselessness of AI ethics. *AI and Ethics*, *3*(3), 869–877. https://doi.org/10.1007/s43681-022-00209-w

Nakpodia, F., Adegbite, E., Olan, F., & Ashiru, F. (2025a). Emergent ethnicity logic and corporate governance in Nigerian boardrooms. *Journal of Accounting and Public Policy*, *54*, Article 107380. https://doi.org/10.1016/j.jaccpubpol.2025.107380

Nakpodia, F., Ashiru, F., Adegbite, E., & Koporcic, N. (2025b). Insidious institutional challenges of mature MNE subsidiaries operating in weak institutional markets: Corporate governance to the rescue. *Journal of International Management*, *31*(4), Article 101258. https://doi.org/10.1016/j.intman.2025.101258

Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society*, *39*(4), 1871–1882. https://doi.org/10.1007/s00146-023-01635-y

Oimann, A.-K., & Tollon, F. (2025). Responsibility gaps and technology: Old wine in new bottles? *Journal of Applied Philosophy*, *42*(1), 337–356. https://doi.org/10.1111/japp.12763

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers*, *25*(1), 123–141. https://doi.org/10.1007/s10796-022-10251-y

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), Article 101885. https://doi.org/10.1016/j.jsis.2024.101885

Plantinga, P., Shilongo, K., Mudongo, O., Umubyeyi, A., Gastrow, M., & Razzano, G. (2024). Responsible artificial intelligence in Africa: Towards policy learning. *Data & Policy*, *6*, Article e72. https://doi.org/10.1017/dap.2024.60

Sætra, H. S. (2023). The AI ESG protocol: Evaluating and disclosing the environment, social, and governance implications of artificial intelligence capabilities, assets, and activities. *Sustainable Development*, *31*(2), 1027–1037. https://doi.org/10.1002/sd.2438

Sahiner, B., Chen, W., Samala, R. K., & Petrick, N. (2023). Data drift in medical machine learning: Implications and potential remedies. *British Journal of Radiology*, *96*(1150), Article 20220878. https://doi.org/10.1259/bjr.20220878

Sanderson, C., Douglas, D., Lu, Q., Schleiger, E., Whittle, J., Lacey, J., Newnham, G., Hajkowicz, S., Robinson, C., & Hansen, D. (2023). AI ethics principles in practice: Perspectives of designers and developers. *IEEE Transactions on Technology and Society*, *4*(2), 171–187. https://doi.org/10.1109/TTS.2023.3257303

Schuett, J. (2024). Risk management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, *15*(2), 367–385. https://doi.org/10.1017/err.2023.1

Schuett, J. (2025). Three lines of defense against risks from AI. *AI & Society*, *40*(2), 493–507. https://doi.org/10.1007/s00146-023-01811-0

Schultz, M. D., Conti, L. G., & Seele, P. (2025). Digital ethicswashing: A systematic review and a process-perception-outcome framework. *AI and Ethics*, *5*, 805–818. https://doi.org/10.1007/s43681-024-00430-9

Schultz, M. D., & Seele, P. (2023). Towards AI ethics' institutionalization: Knowledge bridges from business ethics to advance organizational AI ethics. *AI and Ethics*, *3*, 99–111. https://doi.org/10.1007/s43681-022-00150-y

© IIARJournal. All rights reserved.

Taeihagh, A. (2025). Governance of generative AI. *Policy and Society, 44*(1), 1–22. https://doi.org/10.1093/polsoc/puaf001

Zhang, J., Ki, E.-J., & Cai, Q. (2026). Legitimizing AI: Ethical AI in corporate ethics statements. *Corporate Communications: An International Journal, 31*(1), 21–36. https://doi.org/10.1108/CCIJ-07-2025-0173

© IIARJournal. All rights reserved.