Himachal Pradesh National Law University, Shimla (India)
HPNLU JOURNAL OF LAW, BUSINESS AND ECONOMICS (HPNLU JLBE)

JOURNAL ARTICLES *ISSN: 2584-0436* **HPNLU JLBE**

Volume III (2024)

USE OF ARTIFICIAL INTELLIGENCE IN CORPORATE GOVERNANCE: CONTEMPORARY CHALLENGES

Piyush Bharti & Prachi Kumari

CONTENTS

Essay & Comments

USE OF ARTIFICIAL INTELLIGENCE IN CORPORATE GOVERNANCE: CONTEMPORARY CHALLENGES

Piyush Bharti* Prachi Kumari**

Abstract

Artificial Intelligence (AI) in corporate governance presents opportunities and challenges, particularly in India, where concerns about biased AI systems and a lack of transparency are prevalent. For instance, AI-based facial recognition systems and fintech credit scoring methods have faced criticism for unfair practices. The main idea of this article is to examine these issues and propose solutions, such as reducing biases and implementing clear operational rules for AI systems.

The paper explores AI's implications in decision-making, accountability, and ethics in corporate governance. The article discusses governance challenges faced by AI-driven organisations, including the example of OpenAI's leadership dispute and AI's legal and ethical concerns as a voting board member in Hong Kong. The article also highlights the problem of AI misaligning with human or company values, emphasising the need for close monitoring to ensure ethical alignment. The "black-box" nature of AI, where its decisions are often difficult to explain, is also addressed, suggesting that transparency in AI systems is crucial for fostering trust and fairness.

Additionally, the article explores how AI can influence decision-making at the board level, impacting corporate strategies, shareholder interests, and risk management. These technological advancements must be implemented thoughtfully to avoid unintended consequences and ensure that AI contributes positively to the corporate ecosystem. In conclusion, companies should establish clear objectives to manage AI responsibly, ensure transparency in AI processes, and involve stakeholders in monitoring its usage. By focusing on fairness and accountability, companies can maximise AI's benefits while adhering to ethical standards and responsible governance, ensuring long-term success and sustainability.

In conclusion, companies should establish clear objectives to manage AI responsibly, ensure transparency in AI processes, and involve stakeholders in monitoring its usage. Companies can maximise AI's benefits by focusing on fairness and accountability while adhering to ethical standards and responsible governance.

I

Introduction

Incorporating Artificial Intelligence (AI) into corporate governance in India encounters distinct obstacles indicative of the nation's varied socio-economic context and legal framework. Algorithmic bias is a significant concern, as evidenced by biased facial recognition algorithms. A paper conducted in 2018 by MIT Media Lab revealed that commercial AI systems demonstrated elevated error rates when applied to individuals with darker skin tones for face recognition.[1] This finding further amplifies apprehensions around potential biases in law enforcement and surveillance practices.[2] Indian enterprises must prioritise implementing bias mitigation measures and guaranteeing fairness in AI algorithms.

The issues of transparency and explainability present notable difficulties, as exemplified by the contentious nature surrounding credit scoring methods powered by artificial intelligence. Fintech firms in India are progressively dependent on AI algorithms to evaluate creditworthiness and determine lending outcomes. Nevertheless, the lack of openness in credit scoring systems due to the opacity of these algorithms gives rise to issues, which may result in the exclusion of marginalised people from obtaining financial services. To augment openness, Indian regulatory bodies may contemplate implementing mandatory disclosure obligations for decision-making systems driven by AI, thereby granting customers the ability to comprehend the utilisation and assessment of their data.

So, the corporate landscape is anticipated to undergo significant transformations due to the emergence of disruptive technologies, including AI and Distributed Ledger Technology (DLT). Companies incorporate AI and Distributed Ledger

* Assistant Professor (Law), NIMS University, Jaipur, India.
** Assistant Professor (Law), NIMS University, Jaipur, India.

[1] MIT News, P*aper Finds Gender and Skin-Type Bias in Commercial Artificial-Intelligence Systems*, Massachusetts Institute of Technology, 2018. Available at: https://news.mit.edu/2018/paper-finds-gender-skin-type-bias-artificial-intelligence-systems-0212 (last visited 05-Jan.,2025).

[2] Brian Bergstein et al., *This Week in "Say More": The Computer Scientist Who Became an AI Skeptic* , The Boston Globe , 2 November 2023, Available at https://www.bostonglobe.com/2023/11/02/opinion/joy-buolamwini-artificial-intelligence-say-more-podcast (last visited 05-Jan.,2025).

Technology (DLT), or "CorpTech,"[3] into crucial corporate functions. For example, a Hong Kong venture capital company has appointed VITAL, a robot programmed with AI, to its Board as a voting director.[4]

With the integration of AI in the corporate sector, we are attentively observing the extent of the possibilities for transformation. We observe how the organisation, structures, processes, and institutions of economic life, particularly those controlled by businesses, may change in response to AI's emergence. At the same time, it is crucial to consider the handling of AI in corporate governance and business.

This paper explores the multifaceted challenges and opportunities associated with integrating AI into corporate governance, particularly focusing on India's unique socio-economic and legal framework. The primary research questions include: How does the incorporation of AI influence corporate governance structures and decision-making processes? What are the ethical, legal, and societal implications of employing AI, such as algorithmic bias and the opacity of AI decision-making mechanisms? What governance mechanisms can mitigate the risks associated with AI's integration into corporate governance?

The research addresses several critical problems, including the misalignment between AI-driven systems and corporate governance principles, such as transparency, accountability, and fairness. It highlights the risks posed by algorithmic bias, the opacity of AI decisions, and the challenges of ensuring explainability in AI systems. Additionally, the paper underscores the ethical dilemmas, such as the role of AI in law enforcement and surveillance, and the implications of AI as a decision-making entity, as seen in examples like AI voting directors.

The objectives of the research are threefold. First, it evaluates the legal and ethical dimensions of incorporating AI into corporate governance, particularly in light of India's regulatory and socio-economic landscape. Second, it aims to identify effective governance frameworks and bias mitigation strategies to ensure AI's responsible and ethical use in corporate decision-making. Finally, the paper

---

[3] Luca Enriques and Dirk A. Zetzsche, *Corporate Technologies and the Tech Nirvana Fallacy* , University of Luxembourg *(2020)* https://papers.ssrn.com/abstract=3392321(last visited 05-Jan.,2025).

[4]Business Insider, *A Venture Capital Firm Just Named an Algorithm to Its Board of Directors – Here's What It Actually Does*, Business Insider India, 2014. Available at https://www.businessinsider.in/a-venture-capital-firm-just-named-an-algorithm-to-its-board-of-directors-heres-what-it-actually-does/articleshow/35075291.cms (last visited 05-Jan.,2025).

proposes recommendations for enhancing transparency, accountability, and stakeholder trust in AI-driven corporate governance.

This research aspires to contribute to the evolving discourse on integrating disruptive technologies in corporate governance by addressing these questions and objectives, ensuring they align with ethical standards and societal goals.

II

Testing the Limits of Corporate Governance in AI

The conflict inside the boardroom at OpenAI, the firm responsible for ChatGPT, has brought attention to corporate governance's significance in ensuring AI's safety. There is a widespread consensus that AI will disrupt society, prompting governments to develop legislative approaches to manage its social implications.[5]. Meanwhile, private companies are currently engaged in the development of AI, overseen by executives, monitored by boards of directors, and financially supported by investors.[6] Simply put, the most significant technical advancement of our time is regulated by corporate governance, which consists mainly of rules created by private entities. These rules distribute power and resolve conflicts within a firm.

The case of OpenAI's governance conflict vividly illustrates the intricate interplay between corporate governance and the use of AI in corporate decision-making. The conflict, which led to CEO Sam Altman's termination and most directors' subsequent resignation, underscores the significant influence that corporate governance structures wield over AI-driven organisations.[7]

Furthermore, the possibility of using AI as a voting director, as seen in Hong Kong, introduces various legal and ethical challenges. While AI possesses unparalleled analytical capabilities, its opaque decision-making processes and susceptibility to human biases underscore the need for careful oversight and critical evaluation.

---

[5] Janna Anderson & Lee Rainie, *Solutions to address AI's anticipated negative impacts*, Pew Research Center, Available at: https://www.pewresearch.org/internet/2018/12/10/solutions-to-address-ais-anticipated-negative-impacts/ (Last visited 06-Jan.-2025).

[6] Roberto Tallarita, *AI Is Testing the Limits of Corporate Governance*, HARVARD BUSINESS REVIEW, Dec. 2023, Available at https://hbr.org/2023/12/ai-is-testing-the-limits-of-corporate-governance (Last visited 06-Jan.-2025).

[7] Sam Altman, *OpenAI, and the Importance of Corporate Governance by Lawrence*, SSRN, Availbale at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4679613 (Last visited 06-Jan.-2025).

Section 166(3[8]), which mandates directors to exercise due care, skill, and diligence, raises questions about the suitability of AI directors and the extent to which they can fulfil these obligations.

Ultimately, the case of OpenAI underscores the imperative for robust governance frameworks that can effectively navigate the complexities of AI integration while upholding ethical standards and societal objectives. It calls for a nuanced approach that combines AI's capabilities with human judgment and oversight to ensure responsible decision-making and accountability in corporate governance.

The real problem came when AI was appointed voting director, like in Hong Kong. [9] The human director has different levels of understanding, decision-making capabilities, limitations, calibres, etc. Whereas AI has an extended giant level of capabilities of analysing, understanding, decision making, etc., constraints like the logic and rationale of decisions cannot be traced.

The legal issue is that Section 166(3),[10] provides that directors must "do their duties with due and reasonable care, skill and diligence" and "use their judgment."

For an AI to make accurate decisions, the data it provides must be correct. Since humans code and curate this data, it is likely to be influenced by human biases, which can affect the AI's decisions. Consequently, an AI cannot entirely make its own decisions autonomously. However, an AI director can perform their role with "due and reasonable care" by thoroughly evaluating all relevant factors and making informed choices. Despite this, because the logic and rationale behind the AI's decisions can be opaque, we must also apply common sense and critically assess the AI's judgment rather than unquestioningly trusting it as infallible.

III

Legal Challenges

Technological advancement encounters distinct obstacles and prospects when incorporating AI into corporate governance, and legal recognition, responsibility, and regulations for AI work and decisions must be met. Another challenge is that

[8] Companies Act, 2013, S 166(3).
[9] Supra no., 4.
[10] Supra no. 8.

the logic behind AI decisions is often untraceable, and AI lacks common sense[11]. So, convincing the organisation's stakeholders about the decision with the aid of AI is also a challenge under corporate governance. Sometimes, even minor mishandling, small wrong data input, or wrong interpretation of data can cause great dilemmas. Concerns about transparency, bias and equity emerge regarding artificial intelligence algorithms.[12] The lack of transparency in specific AI models can conceal biases, resulting in biased outputs[13].

Indian enterprises are progressively adopting AI to enhance operational efficiency and gain a competitive edge. However, the legal framework of AI in India is still in its early stages.[14], in contrast to more established markets like the United States and the European Union. As per the Union Minister, the AI framework is expected to be publicly released upon the establishment of a new government.[15] According to him, there has been a productive conversation with the IT industry in recent months regarding the improper use of AI and strategies to prevent it.[16] Although the first step has been taken, more development is needed that focuses on corporate governance separately.

---

[11] Blog, *6 Limitations of AI & Why it Won't Quite Take Over In 2023,* ADOCK SOLUTIONS, Available at: https://www.adcocksolutions.com/post/6-limitations-of-ai-why-it-wont-quite-take-over-in-2023 (Last visited 10-Jan. 2025).

[12] Roxana Daneshjou;Mary P SmithMary & SunMary Sun, *Lack of Transparency and Potential Bias in Artificial Intelligence Data Sets and Algorithms: A Scoping Review*, RESEARCH GATE, Available at: https://www.researchgate.net/publication/354787792_Lack_of_Transparency_and_Potential_Bias_in_Artificial_Intelligence_Data_Sets_and_Algorithms_A_Scoping_Review (Last visited 10-Jan. 2025).

[13] Id.

[14] Rajat Sethi; Deborshi Barat; & Rohin Goyal, *Regulating Artificial Intelligence in India: Challenges and Considerations, S&R ASSOCIATES* (Jul. 5, 2023), Available at: https://www.snrlaw.in/regulating-artificial-intelligence-in-india-challenges-and-considerations/ (Last visited 10-Jan. 2025) .

[15] Indo-Asian News Service, *AI Framework To Be Made Public As Soon As New Government Is Formed: Minister,* NDTV.COM, Available at: https://www.ndtv.com/india-news/ashwini-vaishnaw-says-ai-framework-to-be-made-public-as-soon-as-new-government-is-formed-5503845 (Last visited: 10-Jan-2025).

[16] Id.

The Companies Act, 2013,[17] is the principal legislation regulating corporate governance procedures. Sections 166 and 143[18] delineate directors' tasks and obligations regarding decision-making, accountability, and financial reporting. [19]

The Ethics Guidelines for Trustworthy Artificial Intelligence by the European Union emphasise values such as transparency, accountability, and justice[20]. These guidelines are designed to promote the responsible advancement and implementation of AI technology in diverse areas.[21] The Guidelines derive seven principles from its four core pillars: respect for human autonomy, harm prevention, fairness, and explicability.[22] If businesses adopt AI principles, it will significantly influence corporate governance by bringing accountability and ethics.

AI algorithmic judgments for company directors have to rely on the trustworthiness of the data. Training models often have data-reinforcing biases, and algorithms can contain hidden biases, impacting a director's ability to fulfil their responsibilities with "due and reasonable care" under section 166(3). Biased data can lead to unfair or suboptimal decisions, risking reputational damage and legal liabilities. Hidden algorithmic biases complicate accountability and require directors to have a deeper technical understanding and continuous monitoring to ensure fairness and transparency. To mitigate these risks, directors must exercise rigorous due diligence in AI deployment, seek necessary training, and establish robust ethical standards and risk management practices. Clear communication with stakeholders about AI use and bias management is crucial for maintaining trust and ensuring responsible governance. Ultimately, directors must proactively manage AI technologies to align with company values and legal obligations, ensuring their decision-making processes remain fair, transparent, and accountable. So, what is the impact of AI on directors' ability to fulfil their responsibilities with "due and reasonable care" under section 166(3)?[23] and how to regulate it?

---

[17] Companies Act, 2013, No. 18.

[18] Id., S. 166 & S. 143.

[19] Supra note 2.

[20] Eleanore Hickman & Martin Petrin, *Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective*, 22 EUR. BUS. ORGAN. LAW REV. 593 (2021).

[21] Id.

[22] Id.

[23] Supra no 8; $ 166 (3).

The Ethics Guidelines for Trustworthy Artificial Intelligence suggested by the European Union,[24] provide a comprehensive framework to promote AI technology's responsible and ethical advancement in diverse fields, such as corporate governance. The guidelines emphasise essential principles like transparency, accountability, fairness, and human oversight to minimise AI implementation risks while maximising benefits. It advocates for positive corporate governance values, such as a stakeholder-oriented corporate purpose and diversity, non-discrimination, and fairness. However, their broad nature leaves several questions and concerns unresolved. This paper explores the potential impact and significance of the Guidelines on specific corporate law and governance issues. Greater specificity is necessary to align these principles with company law rules and governance practices. The Guidelines offer a valuable foundation for guiding businesses toward establishing trustworthy AI until more stringent legislative measures are introduced.

Standards of reasonable care or rational decision-making become murkier when AI is just a black box. While explainable AI makes it more understandable, the fundamental problems with data and algorithms cannot be wished away in the name of Explainable AI, as it has already been discussed that AI needs more common sense and tracing the logic of AI decisions is often impossible. Explainable artificial intelligence (XAI) refers to a collection of procedures and techniques that enable human users to understand and have confidence in the outcomes and outputs generated by machine learning algorithms.[25]

Adhering to the principles of trustworthy AI, which refers to artificial intelligence systems that are reliable, ethical, and transparent. These systems prioritise safety, fairness, and accountability, ensuring they operate without bias, respect user privacy, and maintain robust security.[26] Trustworthy AI aligns with human values and adheres to established legal and ethical standards. It is crucial in corporate governance to ensure transparency, accountability, and moral integrity. By integrating ethical considerations into their AI development and deployment processes, companies may improve stakeholder trust, reduce legal and reputational concerns, and promote sustainable long-term growth.

Directors have a fiduciary obligation to exercise care and expertise toward the company and its stakeholders, necessitating the diligent, prudent, and competent execution of their responsibilities. In the age of artificial intelligence, directors face

[24] Supra no. 15.

[25] What is Explainable AI (XAI)?, IBM, (Mar. 29, 2023) Available at: https://www.ibm.com/topics/explainable-ai (Last visited:10-Jan-2025).

[26] Supra no.15.

the challenge of managing the intricacies of technological advancements while upholding their fiduciary duties.

Auditing standards mean the standards of auditing or any addendum for companies or classes of companies referred to in sub-section (10) of section 143[27]. The Indian company laws, under section 149[28]Make it clear that only natural people or individuals can be named directors of businesses. Legal personhood is granted to artificial entities by the law to express and conceptualise the relationships involving the natural persons who are part of the artificial entity, the entity itself, and the relationships among the entity and the external world.

Alignment Problem

A significant concern in the realm of AI safety pertains to the "alignment problem,"[29] superintelligent AI systems may possess goals and objectives that are misaligned with the welfare of human beings.[30] Although it may sound like a science-fiction fantasy, AI researchers widely agree that the development of AI at a level comparable to humans is near, and the alignment issue is real.[31]

While it is possible to programme a highly intelligent AI to aim for socially desirable objectives, it is essential to acknowledge that the AI may also choose to pursue damaging instrumental goals while pursuing these ultimate goals. In corporate governance, AI systems might inadvertently pursue harmful instrumental goals—such as biased decision-making or unethical data use—while striving to achieve ultimate objectives like profit maximisation. This misalignment underscores the importance of robust oversight and ethical guidelines to ensure AI actions align with corporate values and societal norms. The issue is in our current need for more knowledge regarding the methods to consistently instruct AI to align its behaviour with human ideals. While it is possible to enumerate numerous human-compatible behaviours, it is essential to note that this list needs to be more comprehensive. The issue of aligning AI systems resembles the fundamental challenge of corporate governance.

Corporate law and governance's primary focus is addressing the managerial "agency problem," mitigating the danger of managers diverging from investor

[27] Companies Act, 2013, S. 143(10).
[28] Companies Act, 2013, S.149 .
[29] Rob Whiteman, *AI Safety: Alignment Is Not Enough*, MEDIUM (Dec. 27, 2023), Available at: https://medium.com/@rob.w.automation/ai-safety-alignment-isnt-enough-187c1b6a64ac (Last visited on 12-Jan-2025).
[30] Supra no. 6.
[31] *Id.*

preferences.[32] While it does not provide a complete solution, it significantly mitigates the issue.

So, in the context of corporate governance and agency theory, AI alignment problems arise when the actions and decisions of AI systems do not align with the goals and interests of a corporation's stakeholders. Corporate governance involves balancing the interests of shareholders, management, customers, and other stakeholders through rules and practices. Agency theory highlights the potential conflicts between principals (shareholders) and agents (executives), where agents might pursue their interests over those of the principals. AI alignment problems can manifest as goal misalignment, where AI systems are taught or instructed to prioritise short-term gains over long-term sustainability, or information asymmetry, where the outputs of AI systems are not fully transparent or understandable, leading to decisions based on incomplete information. Additionally, there are challenges in monitoring and holding AI systems accountable for their choices, which can result in unintended ethical and social consequences, such as discrimination or privacy violations. To mitigate these issues, corporations need to establish clear objectives for AI, ensure transparency and explainability, implement robust oversight mechanisms, and involve stakeholders in the governance process. By addressing these alignment challenges, companies can effectively leverage AI technologies while maintaining alignment with corporate governance principles and minimising agency conflicts.

IV

AI as a Black Box: Use of Explainable AI to Bring Accountability & Transparency

As AI applications are often black boxes[33], it is not easy to understand their workings, Explainable AI is used to make it more understandable.[34] Despite this, there are still fundamental problems with the data used and algorithms that cannot be wished away by Explainable AI (XAI). One can use XAI to understand AI decisions' logic, rationale, and process. Here are some of the most widely recognised definitions of an explanation:

[32] Id.

[33] Elena Dubovitskaya & Annika Buchholz, *The Management and the Advice of (Un)Explainable AI*, 20 EUR. CO. FINANC. LAW REV. 794 (2023).

[34] Supra no. 25.

i. An explanation refers to the attribution of causal responsibility[35].
ii. An explanation is the act of describing something and addressing why this description is accurate[36].
iii. An explanation is a process of finding or creating an ordinary meaning.[37]

These definitions emphasise that explainable AI produces details or reasons that make its functioning clear or easy to understand.

Firstly, XAI faces the inherent complexity of many AI algorithms and intense learning models. [38]It can be challenging to interpret due to their non-linear and opaque nature. Ensuring consistency and reliability in explanations across different AI systems and applications. Interpretations can vary based on factors such as data quality and feature importance. Because of the intricate algorithms involved, modern AI systems cannot easily explain their decisions, which reduces their practical usability.[39]

Despite the challenges of XAI, which revolve around the complexity of AI models and may impact transparency and accountability in decision-making, XAI and its findings can be used to justify a decision by knowing its reasoning, process, and logic.

Duty to Understand and Care

The "duty to understand" is not an isolated obligation; additional organisational responsibilities accompany it.[40] The "duty to understand" in corporate governance signifies that organisations must not only grasp the intricacies of their operations, markets, and regulatory environments but also integrate this understanding with broader responsibilities. In AI and accountability, this duty encompasses ensuring that AI systems are designed, implemented, and monitored with transparency, fairness, and ethical considerations. Organisations must establish robust governance frameworks that mandate regular audits, impact assessments, and stakeholder engagement to anticipate and mitigate risks associated with

[35] J.R. Josephson & S.G. Josephson, *Abductive Inference: Computation, Philosophy, Technology* APA PsycNet (1996), Available at https://psycnet.apa.org/record/1994-98614-000 (Last Visited: 12-Jan-2025)
[36] Id.
[37] Sajid Ali et al., *Explainable Artificial Intelligence (XAI): What We Know and What Is Left to Attain Trustworthy Artificial Intelligence*, 99 INF. FUSION 101805 (2023).
[38] Id.
[39] Id.
[40]Supra no. 33.

AI.[41][42]Accountability mechanisms must be in place to address any adverse outcomes or biases introduced by AI, ensuring that decision-makers are held responsible for the ethical deployment and consequences of AI technologies. Thus, the duty to understand is intertwined with a commitment to proactive oversight, ethical stewardship, and continuous learning, reinforcing the broader organisational responsibility to foster trust and integrity in their AI initiatives. These responsibilities include carefully selecting the appropriate AI application beforehand and adhering to specific organisational duties while using AI within the company. These duties are designed to prevent AI from making erroneous decisions or, at the very least, to detect and rectify any errors.[43] This primarily involves supervision or monitoring duties (regular and occasional checks) and documentation obligations. In some instances, appointing a Chief Information Officer (CIO)[44] is also recommended. It is misleading to refer to this context as involving "ongoing plausibility checks of algorithmic decisions"; what is likely meant is either Explainable AI (XAI) or the identification of obvious, recurring errors in the AI model's overall behaviour[45].

The described ex-post control could enhance the legal security of using AI in a company, particularly when the AI is designed to make autonomous decisions and actions. However, examining more closely what the duty of care can achieve in this context is essential. If the algorithm provides incorrect advice, management must identify the error before acting on it. Retrospective control that detects the error afterwards is of limited benefit in this situation. If management makes a wrong decision based on AI's erroneous advice, they are responsible for the outcome.

This applies to both legal and business decisions. Suppose a company director follows a legally incorrect AI recommendation. In that case, it constitutes a breach of their duty of legality, potentially making them personally liable for any resulting loss or damage to the company. There is a risk that the director might

---

[41] George Benneh Mensah, (*PDF) Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems*, EGRC Africa Regulation Research CONSORTIUM , Available at : https://www.researchgate.net/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Review_of_Bias_Mitigation_Transparency_and_Accountability_in_AI_Systems (Last visited on 13-Jan-2025.).

[42] Ruchi Garg, *AI Governance Framework & Best Practices*, NEXTGEN INVENT CORPORATION, Available at: https://nextgeninvent.com/blogs/ai-governance-framework-best-practices/ (Last visited: 17-Jan-2025).

[43] Supra no. 33.

[44] Id.

[45] Id.

follow the AI's proposal, leading to an unlawful and discriminatory decision. For entrepreneurial decisions, management might be shielded by the business judgment rule if the decision, based on the AI's recommendation, is economically unfavourable.

The field of explainable AI (XAI) focuses on developing methods to make AI algorithms' computations understandable. This is crucial for ensuring the transparency and fairness of AI programs and creating verifiable, reliable, and trustworthy AI systems. However, XAI does not aim to provide transparency at the level of traditional programs; it does not precisely track every detail for every data point. [46]Instead, it aims to identify which features or training data points are most relevant to the result, or how altering the data might change the outcome. Another approach involves creating simplified models with a comprehensible structure that allows for a certain degree of explainability.

So, the black-box issue in AI pertains to the difficulty of explaining how a system arrives at its decisions. Terms like black-box, grey-box, and white-box indicate different levels of transparency in a system's internal workings.

Explainability and interpretability are closely linked.[47] A method is explainable if humans understand its operations, with interpretability being more commonly used in the ML community.

The black-box issue in AI refers to the challenge of explaining how a system arrives at its decisions, with terms like black-box, grey-box, and white-box indicating different levels of system transparency. Explainability involves making the decision-making mechanisms of models explicit, showing how inputs and outputs are related.[48], often through post-hoc methods. Interpretability focuses on understanding a model's internal workings, making its operations comprehensible to humans. Additional XAI criteria include transparency, which generates human-readable explanations for model decisions; fairness, ensuring unbiased decisions by avoiding inherent data biases; robustness, measuring output sensitivity to input changes and ensuring consistent performance; satisfaction, enhancing the usability and utility of the ML system through explainability; stability, providing consistent explanations for similar inputs; and responsibility, incorporating societal values, morals, and ethics to build trustworthy models. XAI approaches demystify black-box models and align with responsible AI by creating transparent systems. Critical concepts in XAI are interconnected, with explainability

[46] Id.
[47] Supra no. 37.
[48] Supra no. 41.

enhancing interpretability and robustness, affecting satisfaction and reliability. Understanding these concepts helps researchers investigate the field effectively.

The black-box issue in AI involves explaining how AI systems arrive at their decisions, which is crucial for corporate governance as it impacts transparency and accountability. Explainable AI (XAI) addresses this by making decision-making mechanisms explicit and understandable, enhancing transparency, a key corporate governance principle. XAI also ensures fairness by identifying and mitigating biases in AI models, aligning with corporate commitments to equality. Robustness in XAI guarantees consistent performance, which is crucial for reliable corporate operations. By improving the usability and stability of AI systems, XAI supports user satisfaction and predictable outcomes, which are essential for stakeholder trust. Furthermore, XAI incorporates societal values and ethics into AI models, ensuring responsible and ethical decision-making, which is fundamental for corporate governance. In summary, integrating XAI principles helps companies maintain transparency, fairness, robustness, usability, stability, and responsibility, aligning AI operations with corporate governance standards and fostering stakeholder trust.

Conclusion

Incorporating AI into corporate governance in India represents both a transformative opportunity and a significant challenge. The multifaceted obstacles, including algorithmic bias, transparency issues, and the alignment problem, highlight the complex interplay of technology and governance within the socio-economic and legal framework of the nation. These challenges underline the need for a proactive approach to integrating AI responsibly into corporate systems.

A significant finding is the issue of algorithmic bias. The 2018 MIT Media Lab paper, which revealed higher error rates for individuals with darker skin tones in facial recognition systems, exemplifies how AI algorithms can perpetuate systemic inequities. This concern becomes particularly critical when AI is utilised in sensitive domains such as law enforcement or credit scoring, as seen with Indian fintech firms. The opacity of AI-driven credit scoring methods raises ethical concerns about excluding marginalised populations from financial access.

Transparency and explainability are central issues in the AI-corporate governance nexus. Integrating Explainable AI (XAI) can address the "black-box" problem, making AI systems' decision-making processes more transparent and interpretable. XAI ensures that algorithms adhere to fairness, robustness, and accountability principles, aligning with corporate governance requirements.

However, the complexity of modern AI systems necessitates continuous oversight, clear documentation, and the active involvement of stakeholders.

The concept of AI as a voting director, as seen in the example of VITAL in Hong Kong, introduces legal and ethical concerns. While AI's analytical capabilities are unparalleled, its inability to exercise human judgment or provide fully traceable decision rationales raises questions about compliance with legal standards like Section 166(3) of corporate law. This highlights the need for frameworks that balance AI's capabilities with human oversight to uphold corporate accountability and ethical standards.

The alignment problem further complicates AI integration in corporate governance. Misaligned objectives can lead AI systems to pursue instrumental goals that conflict with stakeholder welfare or corporate values. Addressing this challenge requires robust mechanisms for defining and monitoring AI objectives, ensuring they align with societal norms and ethical considerations.

Indian corporations must actively address the challenges of integrating AI into their governance structures by adopting a comprehensive strategy. A key focus should be on **implementing bias mitigation measures**, which involve diversifying training datasets and including diverse perspectives during the design and deployment stages to reduce algorithmic bias.

Mandatory disclosure and explainability are essential for fostering transparency and trust in AI-driven processes. Regulatory authorities in India should require mandatory disclosures for AI systems used in decision-making, and corporations must adopt Explainable AI (XAI) to ensure that stakeholders can comprehend and trust the outcomes generated by these systems.

The establishment of **robust governance frameworks** is another critical step. Governance structures should facilitate the integration of AI into corporate decision-making processes while maintaining human oversight. Regular audits, ethical reviews, and designated roles, such as a Chief Information Officer (CIO) to oversee AI implementation, are necessary to achieve this balance.

Corporations should also emphasise **ethical AI development** by embedding ethical principles within AI systems. This approach ensures that AI operations are aligned corporate and societal values, supported by clearly defined objectives and mechanisms for monitoring adherence to these values.

In conclusion, while integrating AI into corporate governance poses challenges, proactive measures that align technological innovation with ethical, legal, and

societal expectations can address these issues. By adopting these suggestions, Indian corporations can effectively leverage AI to enhance governance while ensuring accountability and fairness in their operations.