
WHEN ALGORITHMS SET CORPORATE INCENTIVES: WHO IS RESPONSIBLE FOR THE DECISION BEFORE THE DECISION?

Nitisha Bhardwaj, Independent Legal Researcher

ABSTRACT

Corporate law usually asks who made a decision. That question becomes harder when an AI system has already shaped what the decision-maker sees. A company may use AI to rank transactions, customers, employees, or compliance matters. A human then makes the final choice. But the human may be choosing from a list that the system has already produced.

This essay argues that corporate oversight should sometimes move one step before the final decision. The concern is not simply that AI makes mistakes or that humans always follow algorithms. The concern arises when a company turns an AI ranking into part of its working process, builds incentives around that ranking, receives serious warning signs that the system is affecting behaviour or causing important matters to be missed, and does not respond. In that situation, the formal presence of a human decision-maker should not end the inquiry. This is not a proposal for a new “AI fiduciary duty.” It is a narrower application of existing oversight principles to a new corporate risk. SEBI's treatment of AI used by regulated entities provides a useful Indian example because responsibility remains with the regulated entity even when the technology comes from a third-party provider. The proposed U.S. MIND Act raises a different but related concern: technology can shape human choice before the final choice becomes visible. Corporate law may therefore need to ask not only who made the decision, but who shaped the conditions in which that decision was made.

Keywords: Artificial Intelligence, Corporate Oversight, Algorithmic Decision-Making, Caremark Doctrine, Board Oversight, AI Accountability.

I. The decision before the decision

Corporate law usually looks for the person who made the decision. That makes sense when the person actually begins with the choice. AI can change that starting point. Consider a company that uses an AI system to rank transactions by fraud risk. An employee then decides which transactions to investigate. The employee can technically reject the ranking and remains responsible for the final choice. But if the employee has time to review only twenty transactions out of one hundred, the position of each transaction on the list matters. The system has not made the decision. It has shaped which decisions are likely to be made.¹

This distinction is easy to miss because the system may produce something that looks like information rather than an instruction. A score of 92 does not say “investigate this transaction.” A ranking does not say “ignore the transaction at number 94.” Yet rankings become important when human attention is scarce. The first twenty cases are more likely to receive attention than the last twenty. A system that controls the order in which matters are presented can therefore influence the allocation of human effort without taking formal decision-making power away from the employee. Scholarship on algorithmic scoring has shown how automated predictions can affect access to important opportunities and make the basis for classifications difficult to challenge.²

The same structure appears outside compliance. A company can rank customers by their likelihood of purchasing, employees by their likelihood of leaving, or transactions by their predicted risk. The manager still decides what to do. But the manager is no longer operating on a neutral field. The system has already decided what deserves attention.

That creates the central problem for corporate accountability. The law can identify the person who made the final choice. It is harder to identify the system that shaped the range of choices presented to that person. Yet people cannot examine everything. Once a company allows an AI system to control the distribution of limited attention, the system becomes part of the company's decision process.

The argument here is therefore not that AI has replaced human decision-making. It is almost

¹ See Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 640–47 (2017).

² See Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 3–8 (2014).

the opposite. *AI can matter to corporate accountability precisely because the human remains in the loop.* The human makes the final choice, but the system may have shaped the conditions under which that choice was made.

The question should therefore not stop at who made the decision. It should sometimes also ask: *who shaped what the decision-maker was given to decide?*

II. From prediction to incentive

The problem becomes more serious when the company does more than display the AI output. It begins to reward employees for acting on it.

Suppose a sales department uses an AI system that gives every customer a score from 1 to 100. The score predicts the customer's likelihood of purchasing. Managers tell employees to prioritize high-scoring customers and measure employee performance by sales. The original function of the system was predictive. Its organizational function is now different. It has become a mechanism for allocating employee time.

This matters because a prediction can become an incentive. Once employees know that the company expects them to follow a ranking and rewards them for the resulting performance, the ranking no longer operates merely as advice. It begins to structure behaviour. The concern resembles the problem associated with Goodhart's Law: once a measure becomes a target, behaviour can change in ways that undermine what the measure was originally intended to capture.³

The important point is that the AI system does not have to be inaccurate for this to happen. Assume the system is correct in predicting that Customer A is more likely to buy than Customer B. The company may still create a poor process if it instructs employees to spend nearly all of their time with customers like A. The model has answered one question correctly: who is more likely to buy? The company has answered a different question: who should receive employee attention? The first answer does not automatically justify the second.

The same distinction matters in compliance. An AI system may correctly identify a large

³ Charles Goodhart, *Problems of Monetary Management: The U.K. Experience*, in *PAPERS IN MONETARY ECONOMICS* 1, 5 (Reserve Bank of Australia 1975); Marilyn Strathern, "Improving Ratings": Audit in the British University System, 5 EUR. REV. 305, 308–09 (1997).

proportion of high-risk transactions. If employees are then rewarded for processing cases quickly, the ranking may become a shortcut for deciding what deserves investigation. Employees may spend less time looking for unusual cases that the model does not rank highly. The model's accuracy can therefore coexist with a serious weakness in the company's process.

This is why accuracy alone is not a sufficient measure of responsible use. Accuracy concerns the relationship between a prediction and an outcome. Oversight concerns what the company does with that prediction. Those are different questions.

Existing scholarship on algorithmic accountability supports this distinction. The legal problem is not limited to whether an automated system produces a technically correct result. It also concerns the institutional setting in which the result is generated and used.⁴ Recent work on corporate AI governance similarly treats AI as a governance problem because the effects of AI depend on how corporations integrate these systems into their existing decision and risk structures.⁵

The concern here is therefore not every algorithmic mistake. It is the company's decision to turn a prediction into an operational incentive and then ignore evidence that the resulting behaviour is creating a serious risk.

III. What *Caremark* can and cannot do

Caremark provides a possible legal route for thinking about this problem, but only if its limits are taken seriously.

In *In re Caremark International Inc. Derivative Litigation*, the Delaware Court of Chancery recognized that directors have an obligation to make a good-faith effort to ensure that reasonable information and reporting systems exist within the corporation.⁶ *Stone v. Ritter* later confirmed the importance of oversight within the duty of loyalty and explained that liability requires a failure of good faith rather than ordinary negligence.⁷

⁴ See Kroll et al., *supra* note 1, at 647–56.

⁵ See Kim Chin Jean Gan, Barbara Wilczek-Stronczek & Angeliki Papasava, *Governing the Algorithm: A Conceptual Review of Strategic AI Oversight in Global Corporations*, AI & SOC'Y (2026), <https://doi.org/10.1007/s00146-026-03207-2>.

⁶ *In re Caremark Int'l Inc. Derivative Litig.*, 698 A.2d 959, 970 (Del. Ch. 1996).

⁷ *Stone ex rel. AmSouth Bancorporation v. Ritter*, 911 A.2d 362, 369–70 (Del. 2006).

The demanding nature of the doctrine matters. *Caremark* does not make directors insurers against corporate failure. A corporation can have an oversight system and still suffer a loss. Nor does every failure to discover a problem establish bad faith.

The later cases show why the nature of the risk matters. In *Marchand v. Barnhill*, the Delaware Supreme Court focused on food safety because the risk was central to the company's business. The court was concerned not simply with the fact that contamination occurred but with the absence of a meaningful board-level system for receiving information about a central risk.⁸ In *In re Clovis Oncology, Inc. Derivative Litigation*, the Court of Chancery similarly examined whether directors had failed to respond to information concerning a matter central to the company's operations.⁹ *Hughes v. Hu* provides another example of the court examining whether a board's oversight mechanisms were meaningful rather than merely formal.¹⁰

These cases do not establish an AI duty. They establish something narrower: where an important corporate risk exists, the question can become whether the board had meaningful mechanisms for receiving information about that risk and responding to it.

Now apply that reasoning to AI.

Imagine that an AI ranking system has become central to a company's compliance process. Employees use it to determine which transactions receive attention. Their performance is partly measured by how quickly they process the cases the system identifies. Complaints then show that important transactions are being missed. An internal audit finds that the model is using outdated information. Employees tell management that they rarely challenge the ranking because doing so slows their work and can affect performance.

At that point, the relevant event is not the original algorithmic error. The company may not have known about that error when the system was deployed. The important event occurs later, when the company has information showing that the system is influencing conduct and may be creating a serious business risk.

That difference matters. A claim based on the original error would ask whether the company

⁸ *Marchand v. Barnhill*, 212 A.3d 805, 821–22 (Del. 2019).

⁹ *In re Clovis Oncology, Inc. Derivative Litig.*, No. 2017-0222-JRS, 2019 WL 4850188, at *11–13 (Del. Ch. Oct. 1, 2019).

¹⁰ *Hughes v. Hu*, No. 2019-0112-JTL, 2020 WL 1987029, at *14–16 (Del. Ch. Apr. 27, 2020).

should have chosen a better model. A claim based on oversight would ask what the company did after it had reason to know that an important system was creating a problem.

The second question fits *Caremark* better.

It also shows why the argument should remain narrow. The mere use of AI should not create fiduciary exposure. Nor should an isolated complaint. The system must be important to the company's operations. The company must have meaningful warning signs. The risk must be serious enough to matter to the business. And those with authority to respond must fail to act despite having the ability to do so.

The claim is therefore not that *Caremark* should be expanded to cover “AI.” It is that an AI system can become the subject of an existing oversight duty when the system itself has become a significant source of corporate risk.

IV. The human in the loop: formal control versus real control

The obvious response is that the human still makes the decision.

That response has force. It would be wrong to assume that people automatically follow algorithms. The empirical evidence is mixed. Research on human-AI interactions has found that people do not simply accept every algorithmic recommendation and may selectively follow or reject algorithmic advice.¹¹ That matters because an argument based on human dependence should not be overstated.

But the absence of universal automation bias does not solve the corporate problem. The relevant question is not whether every person follows an algorithm. It is whether the company's structure can make reliance on the algorithm predictable or difficult to resist.

A 2024 experiment by Daniela Sele and Marina Chugunova found that participants followed algorithmic recommendations more closely when they were given the ability to monitor and adjust them. More importantly, participants were less likely to intervene when the algorithm produced the least accurate recommendations. The authors concluded that the human-in-the-

¹¹ See Saar Alon-Barkat & Madalina Busuioc, *Human-AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice*, 33 J. PUB. ADMIN. RSCH. & THEORY 153, 153–69 (2023).

loop structure in their experiment increased reliance on the algorithm while decreasing the accuracy of the resulting decisions.¹²

That finding matters because “human oversight” is not a binary concept. A human who independently evaluates an AI recommendation is different from a worker who processes hundreds of recommendations under time pressure, with performance targets that reward speed and an organizational message that the AI system is reliable. Both systems contain a human. Only one gives the human a meaningful opportunity to question the machine.

The problem is therefore not that human review is useless. It is that the phrase “human in the loop” can hide the structure around the human.

A company may give an employee formal authority to reject a ranking while making disagreement costly in practice. The employee may have too many cases to review. The employee may be rewarded for speed. The employee may rarely see what happens to cases placed at the bottom of the ranking. The employee may also assume that the company has already tested the system.

In such a setting, the human remains formally responsible while the system shapes the practical conditions of the decision.

A board does not need to supervise each employee's choice. It should know something different: why the company uses the system, what part of the business depends on it, what employees are expected to do with its output, and what happens when employees report that the output does not make sense.

Those are corporate questions.

V. The vendor does not disappear the corporate responsibility

The outsourcing problem is closely related.

A company may obtain an AI system from an outside provider. It may therefore argue that it cannot be responsible for a system because it did not design the model, choose the training

¹² Daniela Sele & Marina Chugunova, *Putting a Human in the Loop: Increasing Uptake, but Decreasing Accuracy of Automated Decision-Making*, 19 PLOS ONE, no. 2, e0298037, at 1–14 (2024)

data, or write the code.

That argument confuses two different responsibilities.

The vendor may be responsible for building the system. The company is responsible for deciding whether and how the system becomes part of its business.

SEBI provides a useful Indian example. In 2019, SEBI required market intermediaries to report their use of AI and machine learning applications and systems.¹³ In 2025, SEBI amended the Intermediaries Regulations to address AI and machine learning tools used by regulated entities. Regulation 16C makes clear that a regulated entity remains responsible for the use of AI and ML tools whether the tools are designed internally or procured from a third-party technology service provider.¹⁴

The provision matters because it rejects a simple “we bought it from someone else” answer. The regulated entity remains responsible for how the technology is used within its business. The provision also covers AI and ML applications used for trading, investment strategies, compliance, management, and other business purposes.¹⁵

SEBI's June 2025 consultation paper continues in the same direction by proposing principles around responsible AI/ML use, including governance, accountability, monitoring, and attention to third-party providers.¹⁶ The consultation paper is not corporate fiduciary law and should not be treated as evidence that Indian law already imposes a general AI oversight duty. Its significance is narrower. It shows a regulatory approach in which the entity using the technology remains responsible for how that technology operates within the regulated activity.

That supports an important distinction. There is a difference between responsibility for building a system and responsibility for placing that system inside a company's decision process.

The vendor may decide how the model is built. The company decides whether employees rely on the model. The vendor may produce the score. The company decides whether the score

¹³ Securities & Exchange Board of India, *Reporting for Artificial Intelligence (AI) and Machine Learning (ML) Applications and Systems Offered and Used by Market Intermediaries*, Circular No. SEBI/HO/MIRSD/DOS2/CIR/P/2019/10 (Jan. 4, 2019).

¹⁴ Securities & Exchange Board of India (Intermediaries) (Amendment) Regulations, 2025, reg. 16C (2025).

¹⁵ Id. reg. 16C(1) & Explanation.

¹⁶ Securities & Exchange Board of India, *Consultation Paper on Guidelines for Responsible Usage of AI/ML in Indian Securities Markets* (June 20, 2025).

determines where employees spend their time. The vendor may update the system. The company decides whether it has a process for responding when employees or internal reviewers raise concerns.

Those are corporate choices.

VI. The MIND Act: influence before choice

The proposed U.S. MIND Act raises a related problem from a different direction.

The **Management of Individuals' Neural Data Act of 2025**, introduced as S. 2925, would direct the Federal Trade Commission to conduct a study on the governance of neural data and related data. It was introduced by Senator Chuck Schumer with Senators Maria Cantwell and Edward Markey as cosponsors.¹⁷

The bill's "Sense of Congress" provision states that neural data and related data can be monetized and used to shape individual behaviour, emotional states, and decision-making. It also identifies concerns about the combination of neurotechnology, artificial intelligence systems, wearable devices, digital platforms, and global data infrastructure.¹⁸

The MIND Act is not a corporate oversight law. It should not be treated as one. Its relevance here is conceptual.

The proposed legislation recognizes a problem that is increasingly important in technology law: influence can occur before a traditional decision is visible. A system may affect what a person notices, what information is presented, or which options appear important without formally making the final choice.

The corporate problem does not require neural data. An ordinary workplace AI system can put one transaction at the top of a screen and another at the bottom. It can determine which employee receives an alert and which customer is considered first. It can influence attention without taking away the formal power to choose.

That creates a difference between formal choice and practical choice. The employee still

¹⁷ Management of Individuals' Neural Data Act of 2025, S. 2925, 119th Cong. §§ 1–3 (2025).

¹⁸ *Id.* § 2.

chooses. But the company has structured the environment in which that choice takes place.

The MIND Act therefore should not be cited as proof that directors already have an AI oversight duty. It does not establish such a duty. Its value is that Congress has expressly recognized a concern about technology shaping behaviour and decision-making before a traditional decision becomes visible.¹⁹

That is the same space in which the corporate problem appears. The employee chooses, but the system has already helped shape the choice.

VII. What the rule should not become

The argument has an obvious danger. Once we accept that AI can shape decisions without making them, it becomes tempting to make companies responsible for almost anything an algorithm influences.

That would be a mistake.

First, there should be no rule that every AI error creates director liability. Algorithms are imperfect. Human decision-making is also imperfect. Corporate law should not turn ordinary technological failure into fiduciary liability.

Second, there should be no general duty requiring directors to understand the technical details of every AI system used by the company. Boards cannot become machine-learning laboratories. Their role is to understand the significance of the system to the business and the structures through which serious risks are reported and addressed.

Third, the argument should not depend on proving that an AI system was biased or unlawful. A system can create a serious corporate risk without violating an AI-specific regulation. The concern may simply be that the system is causing employees to overlook important information or that incentives are making employees follow rankings too closely.

Fourth, the argument should not impose liability simply because the company uses a vendor. The vendor issue matters because outsourcing can make it easier for a company to treat an

¹⁹ Id. § 2(1)–(3).

important internal process as someone else's responsibility.

The proposed rule is therefore deliberately narrow. When an AI system becomes important to a company's decision process, the company knows that employees are relying on it in a way that may create a serious business risk, credible warning signs reach those with authority to respond, and the company fails to act in good faith, existing corporate oversight principles should not become irrelevant merely because a human made the final decision.

This formulation keeps the focus on corporate conduct. The legal question is not whether the algorithm was intelligent or technically perfect. It is whether the company responded responsibly once it knew what the algorithm was doing inside the organization.

Keeping the rule narrow is also important because companies need room to use imperfect technology. No algorithm will be right all the time, just as no human decision process will be free from error. The concern begins when the company knows that the system has become influential and has reason to believe that the way it is being used is creating a serious risk.

At that point, ignoring the system is itself a corporate choice.

VIII. Conclusion

The difficult corporate question may arise before the formal decision.

An AI system can rank the cases, determine the order in which they are seen, and influence where an employee spends limited time. The employee can then make the final choice. From a formal perspective, the human decided. From an organizational perspective, the system may have already shaped the decision.

That does not mean the machine should become the legal decision-maker. It means corporate law should not assume that formal decision-making authority tells the whole story.

The strongest case for oversight arises not when an algorithm makes one wrong prediction but when the company has made the algorithm part of an important decision process and then learns that the system is affecting employee behaviour in a serious way. If the company receives credible warnings and has the ability to respond but does not, the problem is no longer simply technological. It is a problem of corporate oversight.

This approach also explains why the human-in-the-loop defence is incomplete. Human review remains important, and the empirical evidence does not show that humans always follow algorithms. But formal human control can coexist with organizational pressure, performance incentives, limited attention, and repeated reliance on automated rankings. The relevant question is therefore not whether a human technically could disagree. It is whether the company has created a process in which meaningful disagreement is possible and serious warnings are acted upon.

SEBI's approach provides a useful Indian example because responsibility remains with the regulated entity even when AI is obtained from a third party. The MIND Act illustrates the problem from another direction by recognizing that technology can affect human behaviour and decision-making before a traditional decision becomes visible.

Corporate law has long asked who made the decision. With AI, it should sometimes ask a second question: *Who shaped what the decision-maker would see?* That is where the decision before the decision begins.