CALL FOR PAPERS & ARTICLES 2025–26

Organised by LaWGiCo (Nivisam Lawgico Innovation Pvt. Ltd.)

Title:

ATTRIBUTING MENS REA TO CORPORATIONS WHEN AI SYSTEMS CAUSE FINANCIAL OR COMPLIANCE BREACHES: TOWARDS A GOVERNANCE-BASED MODEL OF LIABILITY

By:
Adv. Karan Singh Shekhawat

ATTRIBUTING MENS REA TO CORPORATIONS WHEN AI SYSTEMS CAUSE FINANCIAL OR COMPLIANCE BREACHES: TOWARDS A GOVERNANCE-BASED MODEL OF LIABILITY.

Abstract

The rise of artificial intelligence (AI) in financial and compliance functions has destabilised traditional doctrines of corporate criminal liability, which presuppose that culpability is grounded in human intention. Classical models of *mens rea* attribution such as the identification doctrine, vicarious liability and aggregation of corporate knowledge were developed for organisations whose decisions were taken by identifiable human actors. By contrast, contemporary AI systems operate with a degree of autonomy and opacity that complicates the tracing of a "guilty mind" within the corporate structure. Legal systems in the European Union (EU), United States (US), United Kingdom (UK) and India are responding by shifting emphasis from subjective intention to governance, risk-management and oversight obligations. This paper argues that the trajectory of law and policy is towards a governance-based conception of corporate *mens rea* in AI-driven financial and compliance breaches, in which liability is imputed through defective AI governance, inadequate human oversight and systemic failures in risk management rather than through proof of individual intention alone.

1. Introduction

Mens rea the "guilty mind" remains a foundational element of criminal liability in modern legal systems.[1] In most common law jurisdictions, liability traditionally requires proof of both a prohibited act (*actus reus*) and a culpable mental state such as intention, knowledge, recklessness or negligence.[2] In the United States, the Model Penal Code (MPC) formalised this

[1] Criminal Intent / Mens rea, Cornell Legal Information Institute (LII) (accessed 2025).
[2] Mens rea, Encyclopaedia Britannica (accessed 2025).

into four primary levels of culpability, now widely used as a reference point even where not formally adopted.[3]

When the defendant is a corporation, rather than a natural person, these concepts must be adapted. Corporations are artificial legal persons with no biological mind; yet they are capable of committing serious economic, financial and regulatory offences. To bridge this gap, courts have developed doctrinal techniques for attributing *mens rea* to corporate entities, ensuring that they are not insulated from criminal responsibility merely by virtue of their juridical form.

The rapid deployment of AI in credit scoring, trading, anti-money-laundering (AML) surveillance, sanctions screening and automated reporting has exposed the limits of these doctrines. Contemporary systems based on machine learning can generate complex behaviours that were not expressly programmed, and their internal logic may be opaque even to their designers. When such systems cause discriminatory lending, persistent AML failures or inaccurate regulatory reporting, the central question is: whose *mens rea* if anyone's can be said to underpin the corporate breach?

Comparative regulatory developments in the EU, US, UK and India show a trend towards treating AI-related harm as a function of defective governance rather than purely subjective intention. The EU AI Act, US state-level AI statutes, UK financial-sector guidance and India's evolving digital-governance framework all emphasise transparency, human oversight, risk-management and accountability. This paper situates that trend within corporate criminal jurisprudence and proposes a governance-based model for attributing *mens rea* in AI-driven financial and compliance breaches.

[3] Model Penal Code § 2.02, American Law Institute (selected provisions).

2. Mens Rea and Corporate Criminality

In classical criminal law, *mens rea* denotes the mental element required for an offence typically some combination of purpose, knowledge, recklessness or negligence.[4] Standard formulations describe criminal liability as requiring both *actus reus* and *mens rea*, save where the legislature creates strict-liability offences.[5] The MPC's general part codifies this by providing that no person is guilty of an offence unless they act purposely, knowingly, recklessly or negligently with respect to each material element.

Applied to natural persons, *mens rea* is located in the individual's mental state at the time of the act and may be inferred from words, conduct and surrounding circumstances. For corporate defendants, however, there is no single mind. Corporate criminal law therefore relies on attribution doctrines.

The **identification doctrine** treats the acts and mental state of those who constitute the company's "directing mind and will" typically board members or very senior managers as the acts and state of mind of the company itself.[6] The leading English case *Tesco Supermarkets Ltd v Nattrass* established that only those who are the embodiment of the company's controlling mind fall within this doctrine. Indian law has adopted a similar approach: in *Iridium India Telecom Ltd v Motorola Inc*, the Supreme Court held that the criminal intent of those in control of the company's affairs can be imputed to the corporation, which may be prosecuted for offences requiring *mens rea*.[7]

Vicarious liability offers a different route, particularly in regulatory or public-welfare offences. Under this model, the corporation is liable for wrongful acts committed by employees or agents in the course of their employment and, generally, at least partly for the benefit of the

[4] Tesco Supermarkets Ltd v Nattrass [1972] AC 153 (HL).
[5] "Corporate Criminal Liability: The Iridium/Motorola Case", IndiaCorpLaw Blog (19 November 2010).
[6] Standard Chartered Bank v Directorate of Enforcement AIR 2005 SC 2622; (2006) 4 SCC 278.
[7] Iridium India Telecom Ltd v Motorola Inc (2011) 1 SCC 74.

company.[8] In the US, this respondeat superior standard is widely applied to corporate criminal liability; in India, by contrast, vicarious criminal liability of officers typically arises only where statutes expressly so provide. The Supreme Court in *Aneeta Hada v Godfather Travels & Tours* clarified that where a statute (such as the Negotiable Instruments Act) creates such liability, the company must be arraigned as the principal offender before its officers may be proceeded against.[9]

A further technique is the **collective knowledge** or aggregation doctrine, under which courts piece together fragments of knowledge held by different employees to construct a distinct corporate *mens rea*. This doctrine seeks to prevent corporations from evading liability by compartmentalising information so that no single individual possesses the full picture, even though the organisation as a whole operates with culpable knowledge.

Across these models, the unifying theme is that corporate *mens rea* is derivative: it is constructed from the mental states of natural persons within the enterprise rather than attributed to the entity in its own right.

3. AI and the "Accountability Gap"

AI-driven systems, especially those using machine learning techniques such as neural networks, disrupt these attribution models in three principal ways.

First, AI lacks consciousness or moral agency. However sophisticated, an algorithm cannot "intend" to commit fraud, "knowingly" breach AML rules or "recklessly" discriminate in lending decisions.[10] Legal systems are therefore reluctant to ascribe *mens rea* to AI itself; instead, they must locate culpability in the humans who design, deploy or oversee those systems.

[8] Aneeta Hada v Godfather Travels & Tours (P) Ltd (2012) 5 SCC 661.
[9] *Ibid.*
[10] Freeman Law, "The Collective Knowledge Doctrine Generally" (blog, 2023).

Secondly, advanced models often function as "black boxes" whose internal reasoning is not readily interpretable even by their developers. In high-dimensional models trained on vast datasets, individual decision paths may be difficult to reconstruct ex post. That opacity complicates the evidential process by which prosecutors infer knowledge, recklessness or negligence from corporate decision-making patterns

Thirdly, corporations may seek to weaponise this opacity as an "autonomy defence". They may argue that the harmful behaviour emerged from the model's learning process and was not reasonably foreseeable; that no person within the firm intended or knew of the breach; and that reliance on a third-party AI vendor breaks the chain of attribution. If accepted, such defences would undermine the effectiveness of corporate criminal law precisely in those contexts high-volume, AI-mediated financial activity where regulatory enforcement is most critical

This combination of lack of sentience, technological opacity and potential for autonomy-based defences creates an "accountability gap" in AI-mediated financial and compliance breaches. The core challenge is to close that gap without demanding unrealistic levels of foresight or technical mastery from corporate actors.

4. Regulatory Trajectories in the EU, US, UK and India

4.1 European Union

The EU AI Act establishes a horizontal, risk-based framework for AI systems placed on the EU market.[11] It categorises AI into prohibited practices, high-risk systems, limited-risk systems and minimal-risk systems, with progressively lighter obligations as risk decreases. High-risk AI such as systems used for credit scoring, employment decisions and access to essential services is subject to stringent requirements regarding risk management, data governance, technical documentation, transparency, robustness and human oversight. Article 14 specifically

[11] Regulation (EU) 2024/… on Artificial Intelligence (EU AI Act), esp arts 5, 9, 14, 52. See consolidated text via AI Act Service Desk and commentary.

mandates that high-risk AI systems be designed and developed in such a way that they can be effectively overseen by natural persons, with the aim of preventing or minimising risks to health, safety and fundamental rights. Logging obligations ensure that events are recorded to allow traceability of outputs during operation.[12]

In parallel, the revised **Product Liability Directive (PLD)** modernises EU product-liability law to cover software, AI and digital services It extends strict, no-fault liability to manufacturers and certain operators of AI-enabled products, and expressly treats software including AI systems as "products" for these purposes.[13] Non-compliance with mandatory safety or cybersecurity requirements, or failure to provide necessary updates, can constitute a defect. Member States must transpose the new PLD by December 2026, after which it will apply to products placed on the market.

Although the proposed AI Liability Directive was withdrawn, claimants can rely on national tort and contract law, supported by the AI Act and PLD, to pursue civil claims for AI-related harm. Non-compliance with AI Act obligations is likely to serve as strong evidence of fault or defect in such proceedings.

4.2 United States

At the federal level, the US continues to rely on sector-specific regulation and enforcement by agencies such as the Federal Trade Commission and financial regulators, rather than a single comprehensive AI statute.[14] However, there has been an explosion of AI-related legislation at state level, particularly in areas such as consumer protection, employment, insurance and credit California has enacted provisions adding section 1714.46 to the Civil Code, which bars defendants who developed, modified or used AI from invoking the argument that the AI

[12] European Commission, "AI models with systemic risks given pointers on how to comply with EU AI rules" (18 July 2025).
[13] Directive (EU) 2024/2853 on liability for defective products (revised Product Liability Directive), in force 8 December 2024.
[14] California Civil Code § 1714.46 (Autonomous AI defence barred), text and bill materials.

autonomously caused the harm as a defence in civil actions.[48] This provision explicitly targets the autonomy defence and ensures that responsibility remains with human or corporate actors who deploy AI systems.

Colorado's SB 24-205 introduces one of the first comprehensive state AI laws, imposing governance obligations on developers and deployers of "high-risk" AI systems that make consequential decisions in areas including employment, credit, housing and healthcare.[15] Developers and deployers must implement risk-management programmes, conduct impact assessments, maintain documentation and provide transparency and appeal mechanisms to affected individuals, with obligations calibrated by reference to frameworks such as the NIST AI Risk Management Framework.

These measures interact with existing negligence, discrimination, consumer-protection and financial-services statutes. Courts can treat failure to comply with AI-specific governance duties as evidence of negligence or breach of statutory duty, thereby anchoring liability in governance failures rather than in the internal "intent" of AI models.

4.3 United Kingdom

The UK has chosen a principles-based, technology-agnostic regulatory strategy, relying on existing mandates and cross-sectoral principles rather than enacting a dedicated AI act. [16] The Financial Conduct Authority (FCA), Prudential Regulation Authority (PRA) and Bank of England have jointly emphasised that most of the rules needed to govern AI in financial services on governance, operational resilience, fair treatment of customers and market integrity are already in place.[17]

The FCA's published AI approach and speeches emphasise that firms remain responsible for AI use under existing principles for businesses and the Consumer Duty, and that "agency must

[15] Colorado SB 24-205, Consumer Protections for Artificial Intelligence (2024), and commentary.
[16] FCA, "AI: Moving from fear to trust" (speech, 9 November 2022).
[17] FCA/Bank of England, Artificial Intelligence in UK financial services – 2024 (survey report, 2024).

not be attributed to AI systems" so as to avoid removing accountability from firms and their senior managers. The PRA's supervisory statement SS1/23 on model risk management treats AI and machine-learning models as part of broader model risk and expects banks to adopt a strategic approach to model risk management as a distinct risk discipline.[18] Surveys and discussion papers by the Bank of England and FCA confirm the growing use of AI in UK financial services and the regulators' expectation that firms integrate AI risk within existing governance frameworks. [19]

The Senior Managers and Certification Regime enables regulators to attribute responsibility for AI-related risks and controls to identified individuals, reinforcing a governance-based approach in which AI deployment falls under clearly defined management responsibilities.

4.4 India

Indian law has developed a robust doctrine of corporate criminal liability grounded in *mens rea* attribution to the corporation's "alter ego". In *Standard Chartered Bank v Directorate of Enforcement*, the Supreme Court held that a company can be prosecuted for offences requiring *mens rea* even where the prescribed punishment includes mandatory imprisonment; where imprisonment cannot be imposed, courts may impose fines instead. In *Iridium India Telecom Ltd v Motorola Inc*, the Court confirmed that companies may be liable for offences requiring dishonest intention (such as cheating under the Indian Penal Code), by imputing the intent of those in control of the company's affairs to the corporation. In *Aneeta Hada v Godfather Travels & Tours*, the Court further held that, where statutes create vicarious liability for officers, the company must be arraigned as an accused before officers can be proceeded against. India does not yet have a dedicated AI statute, but its digital-governance framework is evolving.[20] The Digital Personal Data Protection Act 2023 establishes principles of consent,

[18] PRA, Supervisory Statement SS1/23, Model risk management principles for banks (17 May 2023).
[19] FCA, "AI and the FCA: our approach" (9 September 2025).
[20] Digital Personal Data Protection Act 2023 (India) and DPDP Rules 2025;

purpose limitation, data minimisation, storage limitation, security safeguards and accountability in relation to personal data processing, backed by significant administrative penalties of up to INR 250 crore for serious non-compliance Draft rules notified in 2025 elaborate on these principles and create a Data Protection Board to enforce them.[21] These constraints will heavily influence AI model training and deployment, particularly where personal data is involved.

The proposed Digital India Act, intended to replace the Information Technology Act 2000, is expected to address high-risk digital technologies, including AI, and to codify duties of care for intermediaries and platforms in relation to AI-generated content, such as deepfakes and misinformation. Amendments and advisories under the existing IT rules already impose due-diligence and takedown obligations in relation to harmful AI-generated content, with potential loss of safe-harbour status for intermediaries who fail to comply.

Indian courts have begun to confront AI-related harms indirectly, particularly in the context of personality rights. High Courts have granted broad *John Doe* injunctions restraining unauthorised AI-generated depictions and voice clones of public figures, recognising a protectable interest in one's digital persona. Judicial observations have also emphasised that tools such as ChatGPT cannot substitute for human reasoning in adjudication, underscoring the insistence on human oversight in high-stakes decision-making.

Against this backdrop, any defence that "the AI did it" is likely to be weak in Indian law. Given the *Iridium* and *Standard Chartered* line of authority, the decision to deploy high-risk AI systems without adequate safeguards can itself be viewed as a culpable act of the company's directing minds, to which resulting harms may be attributed.[22]

[21] PRS India, "Digital Personal Data Protection Bill, 2023 – PRS Legislative Research" (bill summary and penalties).
[22] EY India, "Decoding the Digital Personal Data Protection Act, 2023" (21 November 2025).

5. Emerging Models of Liability for AI-Caused Financial and Compliance Breaches

Scholars and policymakers have proposed several models for adapting liability doctrines to AI-mediated financial and compliance breaches.

One influential approach reframes many AI systems as products for the purposes of product liability law. By treating AI models, software updates and core algorithmic services as products, strict liability can be imposed on manufacturers and certain operators where a defect in design, training data, documentation or cybersecurity causes harm, without requiring proof of negligence. The EU's revised PLD explicitly extends to software and AI, including defects arising after deployment due to updates, cybersecurity failures or machine-learning-driven changes, reflecting this direction.

Another approach expands vicarious and agency-based liability to AI vendors. Courts in some jurisdictions have been willing to treat vendors of algorithmic decision tools as agents of the deploying entity, enabling claimants and regulators to hold both deployers and vendors responsible for discriminatory or non-compliant outcomes. Academic proposals in the financial-regulation context advocate treating external providers of AML, KYC and market-surveillance systems as agents whose failures can ground direct liability, while still holding financial institutions vicariously liable where they benefit from the automation.

A third strand retains negligence as the organising concept but articulates AI-specific duties of care[23]. Under this model, reasonable care in deploying AI in financial and compliance functions would include: documented pre-deployment risk assessments; bias and robustness testing; transparency regarding model limitations; human-in-the-loop oversight; monitoring for drift and performance degradation; and fallback mechanisms when systems fail Non-compliance with such duties particularly where codified in statutes or regulatory guidance such as the AI Act, Colorado's SB 24-205, or supervisory expectations like SS1/23 can ground liability in

[23] CEPS, legal and policy analyses on EU AI Act and AI liability (various, 2023–2025).

negligence or breach of statutory duty. A. F. Sarch, "Collective Knowledge and the Limits of the Expanded Corporate Criminal Liability Regime" (2024) 44 Oxford Journal of Legal Studies 920.[24]

Finally, governance-based or "failure-to-prevent" models adapt concepts from economic-crime reforms. In the UK, for example, failure-to-prevent offences for bribery and tax evasion impose liability on companies that fail to prevent specified misconduct by associated persons, subject to an "adequate procedures" defence. Recent proposals to extend such models more broadly, along with news of reforms broadening corporate liability through senior-manager regimes, signal an increasing willingness to treat poor governance itself as the gravamen of corporate wrongdoing. Similar logics can be applied to AI-driven compliance failures: a firm that fails to maintain adequate AI-governance procedures could commit an offence of failing to prevent algorithmic misconduct.

Across these models, liability is grounded less in the elusive internal "intent" of AI systems and more in the choices of humans and organisations regarding design, deployment, oversight and control.[25]

6. Conclusion

AI-mediated financial and compliance breaches expose a structural tension in corporate criminal jurisprudence. Traditional attribution doctrines assume that human actors make the relevant decisions and that corporate *mens rea* can be inferred from their individual states of mind. In an era of complex, adaptive AI systems, that assumption no longer holds universally. Comparative legal and regulatory responses in the EU, US, UK and India demonstrate an emerging consensus: AI cannot be treated as an independent locus of agency; firms that design, procure and deploy AI systems remain responsible for their operation; and accountability

[24] Comparative note on Indian and US corporate criminal liability: "Corporate Criminal Liability: A Comparative Analysis between India and the USA" (Mondaq, 2024).
[25] Skadden, "UK regulators publish approaches to AI regulation in financial services" (2 May 2024)

should turn on governance, oversight and risk management. Instruments such as the EU AI Act, the revised PLD, US state laws like California Civil Code § 1714.46 and Colorado SB 24-205, UK financial-sector guidance and India's DPDP Act all reflect this trend.

The most promising way forward is to reconceptualise corporate *mens rea* in AI-driven financial and compliance breaches as institutional culpability grounded in governance failure. Under this model, the "guilty mind" of the corporation is embodied not in anthropomorphic metaphors of a single directing mind, but in the structures, incentives and controls through which it chooses to organise technologically mediated power.

REFERENCES

- Criminal Intent / *Mens rea*, Cornell Legal Information Institute (LII) (accessed 2025). Legal Information Institute+1
- *Mens rea*, Encyclopaedia Britannica (accessed 2025). Encyclopedia Britannica
- Model Penal Code § 2.02, American Law Institute (selected provisions). University of Pennsylvania Law School+2www1.law.umkc.edu+2
- *Tesco Supermarkets Ltd v Nattrass* [1972] AC 153 (HL). corkerbinning.com+3Casemine+3JSTOR+3
- "Corporate Criminal Liability: The Iridium/Motorola Case", IndiaCorpLaw Blog (19 November 2010). IndiaCorpLaw+1
- *Standard Chartered Bank v Directorate of Enforcement* AIR 2005 SC 2622; (2006) 4 SCC 278. Indian Kanoon+2journal.lawmantra.co.in+2
- *Iridium India Telecom Ltd v Motorola Inc* (2011) 1 SCC 74. Mondaq+3Indian Kanoon+3ICSI+3
- *Aneeta Hada v Godfather Travels & Tours (P) Ltd* (2012) 5 SCC 661. Sci API+3Indian Kanoon+3judgmentwindow.in+3

- Freeman Law, “The Collective Knowledge Doctrine Generally” (blog, 2023). Freeman Law+1
- Regulation (EU) 2024/… on Artificial Intelligence (EU AI Act), esp arts 5, 9, 14, 52. See consolidated text via AI Act Service Desk and commentary. SSRN+3Artificial Intelligence Act+3AI Act Service Desk+3
- European Commission, “AI models with systemic risks given pointers on how to comply with EU AI rules” (18 July 2025). Reuters
- Directive (EU) 2024/2853 on liability for defective products (revised Product Liability Directive), in force 8 December 2024. Reed Smith+4Internal Market and SMEs+4Lexology+4
- California Civil Code § 1714.46 (Autonomous AI defence barred), text and bill materials. ai-law-center.orrick.com+3Sidley Austin+3PolicyEngage+3
- Colorado SB 24-205, *Consumer Protections for Artificial Intelligence* (2024), and commentary. Axios+5Skadden+5Colorado General Assembly+5
- FCA, “AI: Moving from fear to trust” (speech, 9 November 2022). FCA
- FCA/Bank of England, *Artificial Intelligence in UK financial services – 2024* (survey report, 2024). Bank of England+2FCA+2
- PRA, Supervisory Statement SS1/23, *Model risk management principles for banks* (17 May 2023). Bank of England+1
- FCA, “AI and the FCA: our approach” (9 September 2025). Financial Times+4FCA+4Regulation Tomorrow+4
- Digital Personal Data Protection Act 2023 (India) and DPDP Rules 2025; see MeitY official text. The Times of India+6MeitY+6Press Information Bureau+6
- PRS India, “Digital Personal Data Protection Bill, 2023 – PRS Legislative Research” (bill summary and penalties). PRS Legislative Research

- EY India, “Decoding the Digital Personal Data Protection Act, 2023” (21 November 2025). EY
- A. F. Sarch, “Collective Knowledge and the Limits of the Expanded Corporate Criminal Liability Regime” (2024) 44 *Oxford Journal of Legal Studies* 920. OUP Academic+1
- CEPS, legal and policy analyses on EU AI Act and AI liability (various, 2023–2025). Internal Market and SMEs+2Pinsent Masons+2
- Skadden, “UK regulators publish approaches to AI regulation in financial services” (2 May 2024). Skadden+1
- Freshfields, “AI regulation in financial services: FCA developments and emerging enforcement risks” (2025). Regulation Tomorrow+1
- Comparative note on Indian and US corporate criminal liability: “Corporate Criminal Liability: A Comparative Analysis between India and the USA” (Mondaq, 2024). Mondaq