International Journal of Economic Practices and Theories (IJEPT)
ISSN: 2247-7225
Volume 2026, Issue 1

Ecological Deception and Algorithmic Shielding: Reimagining Corporate Criminal Liability for AI-Driven Environmental Crimes in the Age of Disruption

Dr. Jagrti Khanna Ahooja, Dr. Sonia Mann

Assistant Professor, Vivekananda Institute of Professional Studies-TC

Assistant Professor, Vivekananda Institute of Professional Studies-TC

Abstract

Industrial regulation has long assumed a human decision-maker behind every act of pollution. That assumption is losing its grip. Corporate decisions affecting environmental quality are increasingly delegated to autonomous AI systems, and where the decision-maker is an algorithm, the standard toolkit of environmental criminal law—built around *mens rea,* corporate attribution, and the directing-mind doctrine—finds nothing to hold. This paper examines that failure through the lens of "Algorithmic Shielding": the phenomenon whereby corporations escape liability under the Polluter Pays Principle by pointing to the inscrutable autonomy of the machine. Drawing on Indian environmental jurisprudence and the EU AI Act framework, it diagnoses the governance gap and advances three interlocking reforms—mandatory algorithmic auditing, Accountability by Design, and a statutory rebuttable presumption of *mens rea*—under the heading of Techno-Legal Accountability.

Keywords: Algorithmic Accountability, Polluter Pays Principle, Environmental Law, Corporate Criminal Liability, AI Governance, Socio-Legal Disruption.

Introduction

"The greatest threat to our planet is the belief that someone else will save it."

— Robert Swan

It is one of the grimmer ironies of the present moment that Swan's warning has found its most troubling institutional echo in the boardroom. India's environmental law rests on a statutory triad[1] and on the constitutional commitments of Articles 48A and 51A(g).[2] Both were drafted with a human polluter in view. Neither contemplates the possibility that the entity actually making the critical environmental decision—how much to discharge, how long to run a hazardous process, how to optimise a plant's throughput—is not a person at all, but a machine-learning system operating beyond any single manager's immediate comprehension.

The gap this creates is not merely theoretical. When AI systems optimise industrial processes, they do so within parameters that may include no meaningful environmental constraint.

[1]Environment Protection Act 1986 (India); Air (Prevention and Control of Pollution) Act 1981 (India); Water (Prevention and Control of Pollution) Act 1974 (India).
[2]Constitution of India, art 48A; art 51A(g).

Outputs that breach pollution norms emerge not from malice, and not always from negligence in any traditional sense, but from the logic of an algorithm that was never instructed to treat regulatory thresholds as binding. The corporation's response, when questioned, is ready-made: the machine decided; no one chose to pollute. This paper names that response "Algorithmic Shielding" and argues that it represents a structural exploitation of inherited legal doctrine that Indian law is, at present, ill-equipped to counter. The EU AI Act[3] has at least begun to reckon with algorithmic accountability; Indian environmental law has not yet joined that conversation.

The argument proceeds in three Parts. Part I traces the foundations of corporate environmental criminal liability and the Polluter Pays Principle in Indian law, and identifies where the edifice cracks when the actor is algorithmic. Part II anatomises the Black Box problem and the governance gap it opens. Part III proposes a Techno-Legal Accountability framework to fill it.

Part I: Corporate Criminal Liability and the Polluter Pays Principle in Indian Environmental Law

No principle is more deeply embedded in Indian environmental jurisprudence than the idea that the cost of pollution must be borne by whoever caused it. The Supreme Court gave that idea doctrinal teeth in *Indian Council for Enviro-Legal Action v. Union of India,*[4] holding that the Polluter Pays Principle is a binding legal rule, not a statement of intent, and that affected corporations must bear both the immediate and long-term costs of remediation. Principle 16 of the Rio Declaration[5] confirms the principle's international pedigree. Together, these authorities establish a framework of compelling normative clarity—so long as it is possible to identify the entity that "caused" the harm.

That identification has traditionally been achieved through the directing-mind doctrine: the corporation is guilty when a natural person at its helm harboured the requisite *mens rea.* The approach works tolerably well when a manager authorised a discharge, or when a board resolution sanctioned a non-compliant process. It does not work when the decision emerged from a training dataset and a gradient descent function. An AI system has no mind to direct. It has no knowledge in any legally cognisable sense. It can produce outputs that systematically exceed every permissible emission limit without any human ever forming an intention, or even an awareness, that this would happen.

The consequences are concrete. A corporation deploying AI over its environmental operations occupies a legally comfortable position: responsible enough to profit from efficiency gains, but—under existing doctrine—arguably not criminally responsible when the system's outputs are harmful. The Polluter Pays Principle says polluters must pay; the directing-mind doctrine, unmodified, cannot confirm that an AI-deploying corporation *is* the polluter in any sense the criminal law can use. This gap—and the impunity it enables—is the core problem the paper addresses.

[3]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence [2024] OJ L 1689 (EU AI Act).
[4]Indian Council for Enviro-Legal Action v. Union of India (1996) 3 SCC 212 (Supreme Court of India).
[5]Principle 16, Rio Declaration on Environment and Development, UN Doc A/CONF.151/26/Rev.1 (Vol. I) (1992).

Part II: The Black Box Problem and the Governance Gap

Pasquale's characterisation of the "Black Box" society[6] identifies a systemic informational asymmetry of direct regulatory consequence: the parties whose algorithmic systems generate the most consequential decisions are also the parties with the least legal obligation to explain them. A machine-learning model trained on commercial and operational data may determine that a production configuration generating unlawful levels of industrial discharge is optimal. The chain of deliberate choice that criminal liability requires does not exist—and the corporation knows this.

Enforcement consequently faces a double difficulty. Proving *mens rea* is near-impossible when the system's architecture and training data are proprietary, and when no individual officer can credibly be said to have "known" what the model would produce. Elvy's work on IoT contracting[7] illuminates the broader pattern: where automated systems intermediate between deployer and consequence, the deploying entity absorbs the benefits while the epistemic risk—"we did not know what it would do"—is displaced onto regulators and victims alike. Algorithmic Shielding is, in essence, the environmental version of that displacement.

Tutt's proposal for pharmaceutical-style pre-market approval of algorithms[8] offers one avenue of reform but remains, in India, entirely unenacted. The existing doctrinal resource that comes closest to bridging the gap is the absolute liability rule of *MC Mehta v. Union of India.*[9] By dispensing with fault as a precondition to liability for hazardous enterprises, the Court fashioned a standard that does not need *mens rea*—but its reach is civil, not criminal, and deterrence of the magnitude required by environmental crime demands more than a compensatory remedy.

A further complication is the self-modifying character of learning systems. Calo notes that AI technologies act in ways their designers did not anticipate,[10] and a system compliant at deployment may, after months of adaptive learning, generate wholly non-compliant outputs. Lessig's insight that code is itself a form of law[11] points toward the remedy: if a system's architecture determines what it will do, then the only meaningful regulation is one that acts on the architecture before deployment, not through enforcement after the event. European environmental liability law,[12] operator-based and strict in liability, was drafted before adaptive machine-learning was a practical industrial reality; it lacks the vocabulary to address systems that are, by design, indefinitely mutable.

[6]Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2015) 3–18.
[7]Stacy-Ann Elvy, 'Contracting in the Age of the Internet of Things: Article 2 of the UCC and Beyond' (2016) 44 *Hofstra Law Review* 839, 860–865.
[8]Andrew Tutt, 'An FDA for Algorithms' (2017) 69 *Administrative Law Review* 83, 90–97.
[9]MC Mehta v. Union of India (1987) 1 SCC 395 (Oleum Gas Leak Case) (Supreme Court of India).
[10]Ryan Calo, 'Robotics and the Lessons of Cyberlaw' (2015) 103 *California Law Review* 513, 530–532.
[11]Lawrence Lessig, *Code and Other Laws of Cyberspace* (Basic Books 1999) 6, 86–88.
[12]Directive 2004/35/CE of the European Parliament and of the Council of 21 April 2004 on environmental liability [2004] OJ L 143/56.

Part III: Towards a Framework of Techno-Legal Accountability

The inadequacy surveyed above is structural, which means the response must also be structural. Incremental judicial interpretation of *mens rea* will not close a gap that is built into the architecture of criminal attribution itself. What is needed is a legislated framework of Techno-Legal Accountability, comprising three pillars.

The first is mandatory Algorithmic Auditing. Corporations deploying AI systems in operations with material environmental exposure—discharge, extraction, manufacturing, waste processing—should be required by statute to commission, at defined intervals, an independent technical audit of that system's decision logic, training data, and capacity to generate non-compliant environmental outputs. Audit reports must be disclosed to the relevant pollution control authority. The obligation to audit and disclose is itself criminally enforceable, regardless of whether actual environmental harm has occurred: the duty is one of preventive transparency. This is the logic of mandatory environmental impact assessment, extended to the algorithmic layer of industrial governance.

The second pillar is Accountability by Design, which draws directly on Lessig's argument[13] that a system's architecture is itself a regulatory instrument. A corporation that builds an AI platform without embedding applicable environmental thresholds as non-negotiable operating constraints has, in effect, designed pollution into its operations. Legislation should therefore require, as a condition of deployment, that AI systems operating in regulated environmental contexts be trained on datasets that incorporate statutory pollution limits, and programmed to treat those limits as constraints that override efficiency maximisation. The EU AI Act's risk-classification framework[14] provides a ready template: AI systems with significant environmental footprints should attract high-risk classification, carrying the most intensive pre-deployment conformity obligations.

The third pillar addresses the *mens rea* deficit through a targeted statutory intervention. Where an AI system deployed by a corporation generates environmental harm, the corporation should be deemed, by statute, to have possessed the requisite *mens rea,* subject to a rebuttable presumption displaceable only on proof that the harmful output was genuinely unforeseeable and that all reasonably practicable precautions were taken. The justification for this reversal is grounded, as Zuboff demonstrates,[15] in the informational advantage that deploying corporations hold over regulators: those who design and train algorithmic systems know far more about their likely behaviours than any enforcement authority can. Allowing that asymmetry to generate criminal immunity is indefensible. The constitutional logic of *MC Mehta* and the absolute liability framework of the National Green Tribunal Act 2010 already establish that those who control the source of risk bear the burden of justifying its consequences.[16]

[13]Lawrence Lessig (n 11) 86–88.

[15]Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019) 63–88.

[16]See generally Usha Ramanathan, 'A Piranha in the Goldfish Bowl: The Right to Information and the Right to Privacy' in B N Kirpal and others (eds), *Supreme but Not Infallible: Essays in Honour of the Supreme Court of*

Way Forward

The Ethics Guidelines for Trustworthy AI issued by the European Commission's High-Level Expert Group[17] insist on accountability, transparency, and human oversight as minimum conditions for legitimate AI deployment. These are not European values; they are governance necessities, and they apply to environmental regulation with at least as much urgency as they apply to banking or healthcare. India, whose industrial AI deployment is accelerating, cannot afford a regulatory posture that treats environmental accountability as a problem to be solved after the ecological damage has been done.

Three legislative priorities are immediate. India's environmental statutes must be amended to expressly address AI-mediated harm, replacing the current silence with a clear attribution rule. The Pollution Control Boards require specialist technical capacity or a dedicated algorithmic audit unit—enforcement authority without technical competence is authority in name only. And India's developing AI governance framework must be drafted in active dialogue with environmental law; responsible AI governance that stops at the factory gate is not responsible governance at all.

The Polluter Pays Principle rests on a moral intuition of compelling simplicity: those who profit from activities that damage the environment must bear the cost of that damage. A corporation that delegates its operational decision-making to an algorithm has not altered that moral calculus—it has merely attempted to alter the legal one. "Accountability by Design" is the legal system's answer to that attempt. It is not a burden on innovation; it is the price of operating a hazardous technology in a society governed by law. The alternative—a regime in which increasing technological sophistication produces decreasing legal accountability—is one that neither the Constitution nor the rule of law can sustain.

India (Oxford University Press 2000); National Environment Tribunal Act 1995 (India); National Green Tribunal Act 2010 (India).

[17]High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' (European Commission, 8 April 2019).