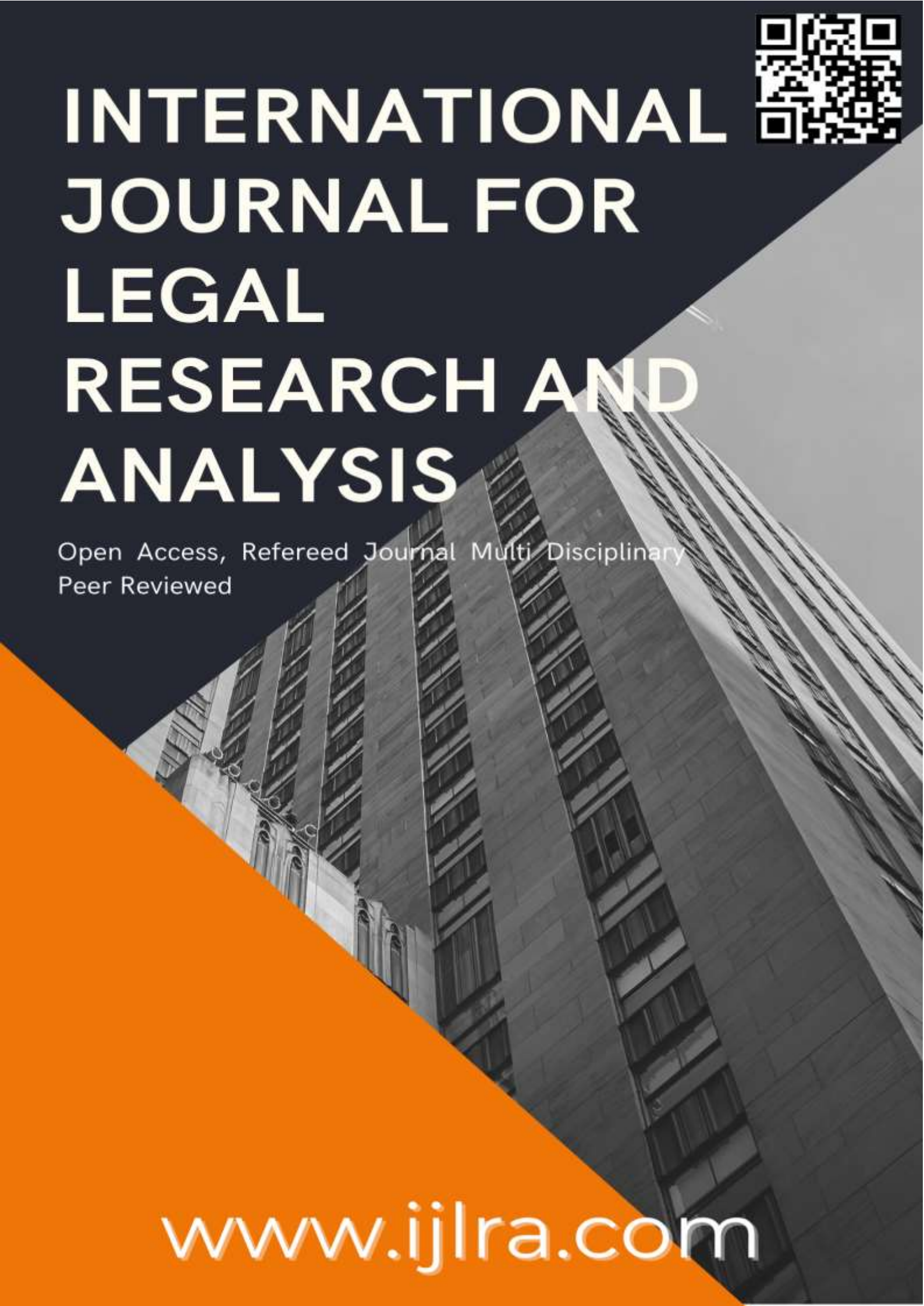
INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed



www.ijlra.com

ISSN: 2582-6433

DISCLAIMER

ISSN: 2582-6433

ABOUT US

The *International Journal for Legal Research & Analysis (IJLRA)* (ISSN: 2582-6433) is a peer-reviewed, academic, online journal published on a monthly basis. The journal aims to provide a comprehensive and interactive platform for the publication of original and high-quality legal research.

IJLRA publishes Short Articles, Long Articles, Research Papers, Case Comments, Book Reviews, Essays, and interdisciplinary studies in the field of law and allied disciplines. The journal seeks to promote critical analysis and informed discourse on contemporary legal, social, and policy issues.

The primary objective of IJLRA is to enhance academic engagement and scholarly dialogue among law students, researchers, academicians, legal professionals, and members of the Bar and Bench. The journal endeavours to establish itself as a credible and widely cited academic publication through the publication of original, well-researched, and analytically sound contributions.

IJLRA welcomes submissions from all branches of law, provided the work is original, unpublished, and submitted in accordance with the prescribed submission guidelines. All manuscripts are subject to a rigorous peer-review process to ensure academic quality, originality, and relevance.

Through its publications, the *International Journal for Legal Research & Analysis* aspires to contribute meaningfully to legal scholarship and the development of law as an instrument of justice and social progress.

ISSN: 2582-6433

PUBLICATION ETHICS, COPYRIGHT & AUTHOR RESPONSIBILITY STATEMENT

The *International Journal for Legal Research and Analysis (IJLRA)* is committed to upholding the highest standards of publication ethics and academic integrity. All manuscripts submitted to the journal must be original, unpublished, and free from plagiarism, data fabrication, falsification, or any form of unethical research or publication practice. Authors are solely responsible for the accuracy, originality, legality, and ethical compliance of their work and must ensure that all sources are properly cited and that necessary permissions for any third-party copyrighted material have been duly obtained prior to submission. Copyright in all published articles vests with IJLRA, unless otherwise expressly stated, and authors grant the journal the irrevocable right to publish, reproduce, distribute, and archive their work in print and electronic formats. The views and opinions expressed in the articles are those of the authors alone and do not reflect the views of the Editors, Editorial Board, Reviewers, or Publisher. IJLRA shall not be liable for any loss, damage, claim, or legal consequence arising from the use, reliance upon, or interpretation of the content published. By submitting a manuscript, the author(s) agree to fully indemnify and hold harmless the journal, its Editor-in-Chief, Editors, Editorial Board, Reviewers, Advisors, Publisher, and Management against any claims, liabilities, or legal proceedings arising out of plagiarism, copyright infringement, defamation, breach of confidentiality, or violation of third-party rights. The journal reserves the absolute right to reject, withdraw, retract, or remove any manuscript or published article in case of ethical or legal violations, without incurring any liability.

ISSN: 2582-6433

ARTIFICIAL INTELLIGENCE AND CRIMINAL LIABILITY: PROSECUTING AI-GENERATED CRIMES AND AUTONOMOUS SURVEILLANCE

AUTHORED BY - LALITH R
Student, School of Law
Vels Institute of Science, Technology & Advanced Studies (VISTAS)

CO-AUTHOR - AKHIL SAJEEV
Assistant Professor, School of Law
Vels Institute of Science, Technology & Advanced Studies (VISTAS)

ABSTRACT

The accelerating integration of artificial intelligence into everyday life has generated profound legal challenges that existing criminal law frameworks are ill-equipped to address. This paper examines the doctrinal and normative gaps that emerge when AI systems autonomously generate crimes or are deployed as instruments of surveillance. By interrogating foundational principles of criminal liability — namely, *mens rea*, *actus reus*, and the principle of individual accountability — the paper demonstrates that traditional attribution models falter when applied to autonomous and semi-autonomous AI actors. Drawing on comparative analysis of Indian, European Union, and common law jurisprudence, alongside emerging regulatory frameworks such as the EU AI Act 2024, this paper proposes a multi-layered liability model that distributes criminal responsibility across developers, deployers, and, where appropriate, corporate entities. The paper further interrogates the legality of autonomous surveillance technologies, evaluating their compatibility with constitutional guarantees of privacy, due process, and the right against self-incrimination. It is argued that the absence of a dedicated AI criminal liability statute in India represents a critical legislative lacuna that demands urgent attention. The paper concludes by advocating for an internationally harmonised regulatory architecture grounded in proportionality, transparency, and the irreducible primacy of human oversight in AI-assisted criminal justice.

Keywords: *Artificial Intelligence, Criminal Liability, Autonomous Surveillance, Mens Rea, AI Regulation, Indian Law, EU AI Act, Machine Learning, Algorithmic Accountability*.

ISSN: 2582-6433

I. INTRODUCTION

Artificial intelligence (AI) is no longer a speculative technology confined to the domain of science fiction. It is embedded in financial systems, healthcare diagnostics, criminal justice prediction tools, and pervasive surveillance infrastructure. As AI systems become increasingly autonomous — capable of perceiving, reasoning, and acting without contemporaneous human direction — they introduce a category of agency that the law has never previously encountered.[1]

Criminal law, by its foundational design, targets human actors who possess the capacity for choice, intent, and moral culpability. The doctrines of *mens rea* and *actus reus* operate on the presupposition that behind every criminal act there exists a mind capable of forming prohibited intent. When an AI system autonomously executes a fraud, a cyberattack, or an act of lethal force, or when it subjects citizens to perpetual algorithmic surveillance, the question of whose mind is responsible — and whether any mind is — becomes legally and morally urgent.[2]

This paper proceeds in five parts. Part II analyses the theoretical challenges AI poses to the classical elements of criminal liability. Part III examines prosecutorial frameworks and liability attribution models, drawing on comparative case law and statutory analysis. Part IV scrutinises autonomous surveillance technologies through the lens of constitutional rights and emerging data protection norms. Part V proposes a legislative and regulatory framework suitable for the Indian context, informed by international best practices. Part VI offers concluding observations.

II. ARTIFICIAL INTELLIGENCE AND THE CLASSICAL ELEMENTS OF CRIMINAL LIABILITY

The architecture of criminal liability rests upon two indispensable pillars: the prohibited act (*actus reus*) and the guilty mind (*mens rea*). These requirements are not mere technicalities; they encode the moral philosophy that punishment is only legitimate where the offender had both the capacity to act otherwise and the culpable state of mind accompanying the wrongful act.[3]

[1]Russell and Norvig, *Artificial Intelligence: A Modern Approach* (4th edn, Pearson 2020) 1–5.
[2]European Parliament, 'Resolution on Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial Authorities in Criminal Matters' (2020/2016(INI)) [2021] OJ C 445/90.
[3]Gabriel Hallevy, *When Robots Kill: Artificial Intelligence Under Criminal Law* (Northeastern University Press

ISSN: 2582-6433

A. The Problem of Mens Rea

Mens rea — the mental element of crime — requires proof that the accused intended the prohibited result, or acted with knowledge, recklessness, or negligence depending upon the offence charged. An AI system, however sophisticated its architecture, does not possess subjective intent in the jurisprudential sense. Its outputs are the product of statistical inference drawn from vast training datasets, calibrated by reward functions designed by its developers.[4] The challenge is compounded by what scholars have termed the 'black box' problem: the internal computations of deep learning systems are often opaque even to their creators, making it practically impossible to reconstruct the 'reasoning' that produced a particular output.[5] If we cannot ascertain why an AI system took a particular action, it is correspondingly difficult to attribute to it — or to any human principal — the requisite mental state for criminal conviction. This epistemic opacity disrupts the evidentiary foundations of criminal prosecution.

B. The Actus Reus Dimension

While the *actus reus* requirement is superficially less problematic — AI systems plainly *act* in the physical world — the voluntariness requirement embedded within the act doctrine raises further complications. Traditional doctrine requires that the act be voluntary, ie, directed by the conscious will of a person. An AI system's 'act' is the product of an automated computational process rather than any conscious volitional choice.[6] The law must therefore determine whether the voluntary act requirement is satisfied by the human decision to deploy the AI, or must be located in each discrete output the system generates.

C. Levels of AI Autonomy and Legal Implications

Gabriel Hallevy's influential taxonomy identifies three paradigms for AI-related criminal liability: perpetration-by-another, natural-probable-consequence, and direct liability.[7] Under the perpetration-by-another model, the developer or deployer who programmes an AI to commit an offence is treated as the principal offender, the machine serving as an innocent instrument — analogous to a person who uses a child or an animal to perpetrate a crime. This model works tolerably well where the harmful behaviour was deliberately designed into the

2013) 23–47.
[4]Lawrence Lessig, *Code: And Other Laws of Cyberspace* (Version 2.0, Basic Books 2006) 120–135.
[5]Ryan Calo, 'Robotics and the Lessons of Cyberlaw' (2015) 103(3) California Law Review 513, 527.
[6]Hallevy (n 3) 57–62; see also Ugo Pagallo, *The Laws of Robots: Crimes, Contracts, and Torts* (Springer 2013) 78–90.
[7]Hallevy (n 3) 89–114.

ISSN: 2582-6433

system.

The natural-probable-consequence doctrine addresses situations where an AI produces harmful outcomes that, whilst not specifically intended, were foreseeable consequences of its deployment. Liability here attaches to those who deployed a system knowing its capabilities but failing to prevent its foreseeable misuse.[8] Direct liability, Hallevy's most controversial category, contemplates treating sufficiently autonomous AI systems as legal persons capable of bearing criminal responsibility in their own right — a proposition that most current legal systems are unprepared to accept, both doctrinally and institutionally.[9]

III. PROSECUTORIAL FRAMEWORKS AND LIABILITY ATTRIBUTION

The prosecution of AI-generated crimes demands a reconceptualisation of existing liability frameworks. Two principal doctrinal approaches merit consideration: vicarious or corporate liability models adapted to AI deployment, and a distributed liability framework that apportions responsibility across the AI value chain.

A. Corporate and Vicarious Liability Models

Anglo-American jurisprudence has long grappled with the attribution of criminal liability to corporate entities. The strict liability approach adopted in *United States v Dotterweich* and *United States v Park* imposed criminal responsibility on corporate officers for regulatory violations without proof of personal knowledge.[10] English law developed the 'directing mind and will' doctrine in *Lennard's Carrying Co Ltd v Asiatic Petroleum Co Ltd*, later applied in *Tesco Supermarkets Ltd v Nattrass*, to identify the natural persons whose acts and mental states could be attributed to the company.[11]

These frameworks, transplanted to AI contexts, yield partial but incomplete solutions. A company that develops and deploys an AI system that perpetrates financial fraud may be subject to corporate criminal liability if the system's harmful outputs were foreseeable and no adequate risk management procedures were instituted. The Indian Penal Code 1860, under sections 34 and 107–116, provides vicarious and abetment liability doctrines capable, in

[8]ibid 95–102.
[9]Gerben Meynen, *Legal Insanity: Explorations in Psychiatry, Law, and Ethics* (Springer 2016) 41–55.
[10]*United States v Dotterweich 320 US 277 (1943); United States v Park* 421 US 658 (1975).
[11]*Lennard's Carrying Co Ltd v Asiatic Petroleum Co Ltd* [1915] AC 705 (HL); *Tesco Supermarkets Ltd v Nattrass* [1972] AC 153 (HL).

ISSN: 2582-6433

principle, of extension to corporate AI deployers.[12] However, these provisions were drafted without any conception of algorithmic agency and require significant judicial creativity to be applied to AI scenarios.

B. The Distributed Liability Framework

A more theoretically coherent approach distributes liability across the AI value chain — comprising designers, trainers, deployers, and end-users — in proportion to each actor's knowledge, control, and contribution to the harmful outcome.[13] This framework draws on the principles underlying product liability in tort law but adapts them to the criminal context by incorporating the mental element requirements appropriate to the relevant category of offence. Designers bear primary responsibility where the harmful output was a foreseeable and preventable consequence of architectural choices — for example, training a system on biased datasets, or failing to implement adversarial robustness safeguards. Deployers bear responsibility where they deploy systems in contexts their capabilities render dangerous, or fail to establish adequate human oversight mechanisms. End-users may bear responsibility where they override safety controls or deliberately weaponise AI capabilities.[14]

C. The Information Technology Act and AI-Generated Cybercrimes

India's Information Technology Act 2000 (IT Act), particularly sections 43 and 66, criminalises unauthorised access to computer systems and related offences.[15] As AI increasingly becomes the instrument of choice for cyberattacks — deploying adaptive malware, automated spear-phishing, and adversarial exploits — prosecutorial reliance on these provisions becomes indispensable. However, the IT Act's definitional architecture does not contemplate AI agency; it addresses human-directed computer misuse. Judicial interpretation will be required to determine whether an AI-executed cyberattack, without direct contemporaneous human direction, falls within the statutory prohibition.

The Model Penal Code's graduated *mens rea* framework — purpose, knowledge, recklessness, and negligence[16] — offers a useful analytical template for Indian legislative drafters seeking to calibrate AI criminal liability to the degree of advertence a human principal had to the risk of

[12]Indian Penal Code 1860 (IPC), ss 34, 107–116.
[13]Andrew Selbst and Solon Barocas, 'The Intuitive Appeal of Explainable Machines' (2018) 87 Fordham Law Review 1085, 1098–1110.
[14]Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale University Press 2021) 175–195.
[15]Information Technology Act 2000 (IT Act), ss 43, 66, 66A.
[16]Model Penal Code (American Law Institute, 1962) s 2.02.

ISSN: 2582-6433

AI-generated harm. A developer who knowingly programmes an AI to deceive consumers satisfies the purposive standard; one who ignores well-documented risks of algorithmic manipulation may satisfy the recklessness or negligence threshold.

IV. AUTONOMOUS SURVEILLANCE: LEGALITY, ACCOUNTABILITY, AND CONSTITUTIONAL LIMITS

Autonomous surveillance technologies — encompassing facial recognition systems, predictive policing algorithms, and persistent CCTV analytics — represent a qualitatively distinct challenge from AI-generated crimes. Here, the State itself is the deployer of AI power, and the individuals subjected to surveillance are presumptively innocent citizens exercising constitutionally protected rights. The legal issues therefore engage not only criminal law but constitutional and human rights law.[17]

A. Facial Recognition and the Right to Privacy

In the landmark judgment of *K S Puttaswamy v Union of India* (Puttaswamy I), the Supreme Court of India unanimously held that the right to privacy is a fundamental right under Article 21 of the Constitution.[18] The Court recognised that informational privacy — the right to control the collection and use of personal data — is integral to human dignity and autonomy. Autonomous facial recognition surveillance, by its nature, systematically captures biometric data of individuals without their knowledge or consent, generating searchable profiles of their movements, associations, and behaviours. Such systems represent a profound interference with the right recognised in *Puttaswamy I*, and must satisfy the constitutional tripartite test of legality, necessity, and proportionality.

The English Court of Appeal's decision in *R (on the application of Edward Bridges) v Chief Constable of South Wales Police* is instructive in this regard. The Court held that the deployment of automated facial recognition technology by the police was unlawful, as the relevant legal framework was insufficiently clear to satisfy the requirement of lawful basis for interference with Article 8 of the European Convention on Human Rights.[19] Indian courts, applying equivalent proportionality analysis, would be required to evaluate whether existing statutory powers — principally under the IT Act and the Code of Criminal Procedure 1973 —

[18]*K S Puttaswamy v Union of India* (2017) 10 SCC 1 (Puttaswamy I); *K S Puttaswamy v Union of India* (2019) 1 SCC 1 (Puttaswamy II – Aadhaar).
[19]*State v Loomis* 881 NW 2d 749 (Wis 2016); *R (on the application of Edward Bridges) v Chief Constable of South Wales Police* [2020] EWCA Civ 1058.

ISSN: 2582-6433

provide a sufficiently precise legal basis for mass automated biometric surveillance.

B. Predictive Policing and Algorithmic Discrimination

Predictive policing algorithms purport to forecast criminal activity by analysing historical crime data, socio-economic indicators, and demographic profiles. These systems, however, are demonstrably susceptible to encoding and amplifying the structural biases embedded in the datasets upon which they are trained.[20] Where historical data reflects discriminatory policing practices — for example, the over-policing of marginalised communities — the algorithm will systematically over-predict criminality in those communities, generating a self-reinforcing cycle of discriminatory surveillance and enforcement.

The Wisconsin Supreme Court's decision in *State v Loomis* raised profound concerns about the use of the COMPAS recidivism prediction algorithm in sentencing decisions, noting that the algorithm's opacity precluded meaningful judicial review of its outputs.[21] This opacity problem is equally acute in the policing context: where an individual is subjected to enhanced surveillance or preventive detention on the basis of an algorithmic risk score, their right to challenge the factual and inferential basis of that determination — the right to a fair hearing guaranteed by Article 21 of the Indian Constitution — is substantially compromised.

C. Autonomous Weapons and International Humanitarian Law

Lethal autonomous weapon systems (LAWS) represent the most extreme instantiation of autonomous AI acting in the criminal law context. A LAWS capable of selecting and engaging targets without human intervention raises the question of whether any accountable person can be identified when civilian casualties result — the so-called 'accountability gap'.[22] International humanitarian law requires that targeting decisions comply with the principles of distinction, proportionality, and precaution.[23] These principles presuppose a human decision-maker capable of contextual moral judgment — a presupposition that fully autonomous lethal systems cannot satisfy.[24]

---

[20]Marjorie Zatz and Nancy Rodriguez, *Punishing Immigrants: Policy, Politics, and Injustice* (New York University Press 2015) 133–148; see also Julia Angwin and others, 'Machine Bias' (ProPublica, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 10 January 2025.

[22]Suranga Seneviratne, 'Autonomous Weapons and the Limits of Analogy' (2019) 21 Journal of Military Ethics 48, 56–63.

[23]Additional Protocol I to the Geneva Conventions 1977, Art 36; Rome Statute of the International Criminal Court 1998, Art 8.

[24]Human Rights Watch and International Human Rights Clinic, 'Losing Humanity: The Case Against Killer Robots' (November 2012) 3–11 <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots> accessed 12 January 2025.

ISSN: 2582-6433

The Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts has been engaged in ongoing deliberations regarding LAWS since 2014, but has thus far failed to produce a binding instrument.[25] The accountability gap LAWS create — where neither the machine (having no legal personality), nor the commander (lacking sufficient human control over the targeting decision), nor the developer (lacking the requisite criminal intent) can be straightforwardly prosecuted — represents the most acute manifestation of the broader problem this paper addresses.

V. TOWARD A LEGISLATIVE FRAMEWORK: INDIA IN COMPARATIVE PERSPECTIVE

A. The EU AI Act 2024: A Regulatory Template

The European Union AI Act 2024 constitutes the world's first comprehensive legal framework governing AI across all sectors.[26] Its risk-based architecture classifies AI systems into four tiers — unacceptable risk (prohibited), high risk (subject to conformity assessment), limited risk (subject to transparency obligations), and minimal risk (unregulated) — and imposes graduated obligations accordingly.[27] Real-time remote biometric identification systems deployed in public spaces for law enforcement purposes are classified as high-risk and subject to stringent requirements including fundamental rights impact assessments, human oversight mechanisms, and logging of operations.

The GDPR's Article 22, which prohibits solely automated decision-making producing significant legal effects without adequate human review,[28] provides an important cognate principle that Indian data protection law should incorporate by reference. The Digital Personal Data Protection Act 2023 currently lacks an equivalent provision directly applicable to criminal justice contexts.[29]

B. India's Legislative Lacuna and the Way Forward

India's existing legal architecture — the IPC 1860, the IT Act 2000, the DPDP Act 2023, and

[25]Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons (CCW) 1980; Report of the Group of Governmental Experts on Lethal Autonomous Weapons Systems, UN Doc CCW/GGE.1/2019/3 (25 September 2019).
[26]European Parliament and Council Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence [2024] OJ L 1689/1 (EU AI Act).
[27]ibid, Arts 6–51 (risk classification); Art 9 (risk management systems).
[28]Regulation (EU) 2016/679 (General Data Protection Regulation) [2016] OJ L 119/1, Arts 13–14, 22.
[29]Personal Data Protection Bill 2019 (India), cls 25–32; Digital Personal Data Protection Act 2023 (India), ss 7–14.

ISSN: 2582-6433

the Bharatiya Nyaya Sanhita 2023 — collectively fail to address the distinct challenges of AI criminal liability and autonomous surveillance.[30] The NITI Aayog's National AI Strategy, while forward-looking, is a policy document without legislative force.[31] A dedicated AI Governance and Liability Act is urgently required, incorporating the following core elements. First, the statute should establish a definition of 'autonomous AI system' calibrated to the degree of independence from contemporaneous human direction, recognising that liability attribution must be sensitive to the position on this spectrum that a given system occupies. Second, a duty of care should be imposed on all developers and deployers of AI systems operating in high-risk sectors — criminal justice, defence, healthcare, financial services — requiring prior conformity assessment, algorithmic impact assessment, and continuous monitoring. Third, a distributed liability model should be codified, allocating criminal and civil liability among developers, deployers, and users in accordance with their respective knowledge, control, and contribution to foreseeable harm.[32]

Fourth, the statute should establish a mandatory human oversight requirement for AI systems deployed in criminal justice and law enforcement — prohibiting purely automated decisions that produce adverse legal consequences for individuals without meaningful human review. This requirement, modelled on GDPR Article 22 and the EU AI Act, directly addresses the due process concerns identified in *Bridges* and *Loomis*. Fifth, the statute should codify a prohibition on deployments of AI systems that perpetuate structural discrimination, operationalising the fundamental right to equality under Articles 14 and 15 of the Constitution of India in the algorithmic context.[33]

C. International Harmonisation and the Role of Multilateral Bodies

The inherently transboundary character of AI development and deployment renders unilateral national legislation insufficient. AI systems trained in one jurisdiction are routinely deployed in others; criminal actors leverage AI infrastructure hosted across multiple states; and LAWS may be deployed by one State in the territory of another.[34] A coherent international response requires harmonised minimum standards — analogous to those established by the Financial

[30]Pavan Duggal, 'Artificial Intelligence and the Law in India: An Emerging Paradigm' (2020) 1 Indian Journal of Artificial Intelligence and Law 1, 8–14.
[31]National Institution for Transforming India (NITI Aayog), *National Strategy for Artificial Intelligence* (June 2018) 65–78.
[32]Hallevy (n 3) 147–162; Pagallo (n 6) 145–160.
[33]Digital Personal Data Protection Act 2023 (India), s 4; Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules 2011, r 5.
[34]Calo (n 5) 533–538; see also Jack Balkin, 'The Path of Robotics Law' (2015) 6 California Law Review Circuit 45, 50–58.

ISSN: 2582-6433

Action Task Force in the anti-money laundering context — accompanied by mutual legal assistance arrangements specifically adapted to AI-generated transnational crime.

The OECD Principles on Artificial Intelligence and the United Nations Secretary-General's Roadmap for Digital Cooperation provide important political commitments but lack binding legal force.[35] India should actively engage in multilateral negotiations toward a binding international instrument on AI liability, bringing its perspective as the world's largest democracy and a leading AI development nation. The proposed instrument should address, at minimum: common definitions of AI agency and autonomous systems; minimum standards for AI safety assessment and conformity certification; protocols for mutual legal assistance in AI-facilitated transnational crime; and a binding prohibition on LAWS that operate without meaningful human oversight of targeting decisions.[36]

VI. CONCLUSION

The question of AI criminal liability is not a peripheral curiosity of jurisprudence; it is a foundational challenge for legal order in the twenty-first century. As AI systems assume greater autonomy in perpetrating frauds, launching cyberattacks, making targeting decisions in armed conflict, and governing the surveillance of citizens, the law's insistence on human agency as the locus of criminal responsibility is placed under sustained and escalating strain.

This paper has argued that existing criminal law doctrines — the *mens rea* and *actus reus* requirements, corporate liability attribution, and the directing mind doctrine — provide partial but ultimately insufficient resources for the prosecution of AI-generated crime. The distributed liability framework proposed herein — allocating responsibility across developers, deployers, and users in proportion to knowledge and control — represents the most defensible adaptation of existing principles to the AI context, pending more comprehensive legislative intervention.

With respect to autonomous surveillance, the paper has demonstrated that facial recognition, predictive policing, and lethal autonomous weapons raise acute constitutional concerns — engaging the fundamental rights to privacy, equality, and fair process recognised by the Supreme Court in *K S Puttaswamy v Union of India*. Existing Indian legislation provides no adequate regulatory framework for these technologies; the *Bridges* decision and the EU AI Act

[35]Recommendation of the OECD Council on Artificial Intelligence, OECD/LEGAL/0449 (22 May 2019); United Nations Secretary-General's Roadmap for Digital Cooperation (June 2020) 34–38.

[36]Asaro Peter, 'On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making' (2012) 94(886) International Review of the Red Cross 687, 705.

ISSN: 2582-6433

2024 offer instructive comparative models.

The paper has argued for a dedicated AI Governance and Liability Act incorporating risk-based classification, mandatory human oversight requirements, conformity assessment obligations, and a codified distributed liability framework. At the international level, India should lead engagement toward a binding multilateral instrument establishing minimum standards for AI liability and a prohibition on fully autonomous lethal systems.[37]

The foundational commitment of criminal law — that the power to punish is legitimised by the capacity of individuals to choose and to know — must not be abandoned in the face of algorithmic complexity. It must instead be reinvented: through careful institutional design, proportionate regulation, and an unflinching insistence that the benefits and burdens of AI are distributed according to the demands of justice.

[37]Hallevy (n 3) 170–185; see also Tom Allen and Robin Widdison, 'Can Computers Make Contracts?' (1996) 9 Harvard Journal of Law and Technology 25, 45.