IJMRRS

International Journal for Multidisciplinary
Research, Review and Studies

ISSN: 3049-124X (Online)

VOLUME 2 - ISSUE 1

2024

Balancing Technology and Justice: The Imperative for Algorithmic Accountability in India's Criminal Justice System

Author: Chandrika Singh Student of LL.M in Criminal Law from Amity University, Lucknow.
Co-author: Dr. Kavya Chandel Assistant Professor at Amity University, Lucknow.

Abstract

The increasing integration of artificial intelligence (AI) into India's criminal justice system represents a revolutionary shift, bringing both remarkable efficiencies and significant risks. While AI promises to enhance decision-making through predictive analytics, automated case management, and surveillance tools, it also presents a critical challenge: accountability. In a legal system that has traditionally been grounded in human intention, how does one hold an algorithm accountable when it amplifies bias, obscures reasoning, or nudges a system towards unjust outcomes? This dilemma underscores the need for a robust framework of algorithmic accountability in India's criminal law. This article explores the social and legal implications of AI in criminal justice, focusing on the balance between technology and justice, the accountability of AI systems, and the reforms necessary to navigate the intersection of law and machine learning. Drawing on India's constitutional jurisprudence, emerging data protection legislation, and comparative insights from the European Union's AI Act, this article argues that algorithmic accountability must be treated as a socio-legal imperative embedded in enforceable legal standards rather than left to aspirational policy.

Keywords: Algorithmic Accountability, Artificial Intelligence, Criminal Justice, Predictive Policing, Constitutional Law, Data Protection

I. Introduction: Policing the Black Box

The rise of algorithmic systems within India's criminal justice framework is inevitable, with courts and police increasingly relying on AI for a variety of purposes. From research tools to transcription and case management aids, AI is already a fixture in many legal processes. One of the most concerning applications is predictive policing: AI systems used to forecast criminal activity based on historical data. While such tools claim to increase efficiency, they also carry inherent risks: algorithms can perpetuate systemic biases, obscure the reasoning behind their decisions, and lead to wrongful arrests.

When these systems fail or cause harm, the question must be asked: who is responsible? Is it the coder who wrote the algorithm, the law enforcement officer who relied on it, the tech company that developed the tool, or the state that deployed it? AI in criminal justice introduces complexities that traditional criminal law was not designed to address. Most notably, the assumption that crime is rooted in human intention or negligence creates a significant gap when AI systems are involved.

Criminal law in India is built on the foundation of mens rea (the guilty mind), where intent is presumed to be a necessary condition for criminal liability. However, algorithms do not "intend", they simply generate outputs. The real question is whether anyone can be held accountable when AI, through human decisions, creates unjust outcomes. This article argues that the answer must be yes, but only if India's legal framework is reformed to locate and assign accountability across the chain of human decisions that produce, deploy, and rely upon algorithmic systems.

Research Problem

The central research problem addressed in this study revolves around the inadequacy and fragmentation of existing legal and policy frameworks in effectively regulating AI-driven criminal justice systems. While several countries have introduced guidelines and ethical principles for AI governance, there remains a lack of uniformity and enforceability in these frameworks. In the Indian context, regulatory efforts are still in a nascent stage, with multiple authorities operating in silos. This fragmented approach creates ambiguity regarding jurisdiction, compliance requirements, and liability in cases of harm caused by AI systems. The rationale for this research lies in the urgent need to critically examine the existing regulatory landscape, identify gaps and inconsistencies, and propose a cohesive governance model that addresses the unique challenges posed by AI in criminal justice.

Research Objective

The primary objective of this study is to analyze the current policy and regulatory frameworks governing AI in the criminal justice sector and assess their effectiveness in addressing emerging challenges. The study aims to evaluate the extent to which existing laws and guidelines ensure accountability, transparency, data protection, and ethical compliance in AI-driven criminal justice systems. Additionally, it seeks to explore comparative international approaches to AI governance and identify best practices that can be adapted to the Indian context. Another key objective is to propose policy recommendations that can contribute to the development of a comprehensive and coherent regulatory framework for AI in criminal justice.

Research Questions

To achieve these objectives, the study is guided by several key research questions. These include:

1. What are the existing legal and policy frameworks governing AI in criminal justice, both in India and globally?
2. To what extent do these frameworks address issues of accountability, liability, data privacy, and ethical concerns?
3. What are the major gaps and challenges in the current regulatory landscape?
4. How have other jurisdictions approached AI governance in criminal justice, and what lessons can be drawn from their experiences?

5. What policy measures can be implemented to ensure effective and responsible use of AI in the criminal justice sector?

Scope of Study

The scope of this study is primarily confined to the governance and regulatory aspects of AI in criminal justice, with a focus on legal, ethical, and policy dimensions. It examines the application of AI technologies within policing, prosecution, adjudication, and case management settings. The study does not delve deeply into the technical functioning of AI algorithms, as its primary emphasis is on legal and policy analysis. Geographically, while the research is centered on India, it incorporates a comparative perspective by examining regulatory approaches in jurisdictions such as the European Union. The study is subject to certain limitations, including the rapidly evolving nature of AI technology and the lack of extensive empirical data on the real-world impact of AI in criminal justice proceedings.

Research Methodology

The research methodology adopted in this study is primarily doctrinal in nature, involving a detailed analysis of existing legal statutes, policy documents, regulatory guidelines, and judicial pronouncements relevant to AI in criminal justice. This includes the examination of national laws, international frameworks, and constitutional jurisprudence governing data protection, criminal procedure, and technology regulation. In addition to doctrinal analysis, the study incorporates a limited non-doctrinal approach by referring to secondary sources such as academic literature, policy reports, and case studies to provide contextual understanding and support the analytical framework.

II. Conceptual Framework of AI in the Criminal Justice System

Artificial Intelligence in the criminal justice context refers to computational systems capable of performing tasks that traditionally require human judgment, such as assessing risk, identifying patterns in criminal behaviour, or processing evidentiary materials. These systems encompass machine learning algorithms that learn from historical data, natural language processing tools used to analyze legal documents and court records, computer vision systems deployed in facial recognition and surveillance, and predictive analytics platforms used to forecast criminal activity or recidivism risk.

India's Bharatiya Nyaya Sanhita (BNS), which replaced the Indian Penal Code on July 1, 2024, carries forward a core assumption: crime is anchored in human intention, knowledge, or negligence. Under the BNS, the term "person" includes not only natural persons but also companies and associations, thus allowing for corporate criminal liability. However, AI systems operate in a manner that bypasses traditional understandings of human intention and decision-making. An AI system can make a prediction or decision that directly affects an individual's liberty, but if the result is biased or erroneous, accountability cannot rest with the machine itself, there is a structural gap between the output of the machine and human responsibility.

The deployment of AI in criminal justice is multifaceted. Predictive policing tools claim to forecast criminal activity based on historical data, enabling pre-emptive law enforcement. Facial recognition systems are used at public gatherings, border crossings, and in investigative proceedings. Automated case management systems reshape how courts process filings, schedule hearings, and manage evidence. AI-assisted transcription and legal research tools are increasingly embedded in courtroom processes. Each of these applications raises distinct legal and ethical questions that existing frameworks are ill-equipped to resolve.

Despite its potential utility, the deployment of AI in criminal justice raises significant ethical and legal concerns. Algorithmic systems trained on biased historical data may perpetuate and amplify existing patterns of discrimination. The opacity of many AI models—the "black box" problem, undermines the ability of affected individuals to understand, challenge, or contest automated decisions. Questions of accountability, liability, and procedural fairness remain unresolved within India's current legal architecture, necessitating a comprehensive conceptual and regulatory response.

III. Need for Policy Framework and Governance

The rapid integration of artificial intelligence into the criminal justice sector has created unprecedented opportunities for efficiency and enhanced law enforcement capability, but it has simultaneously underscored the urgent need for a robust policy framework and governance structure. In the absence of comprehensive regulation, the deployment of AI systems in policing, prosecution, and adjudication poses significant risks: particularly in the form of algorithmic bias, inaccuracies in decision-making, and potential violations of fundamental rights.

AI systems trained on incomplete or non-representative datasets may produce biased outcomes, disproportionately affecting marginalized or underrepresented populations and thereby reinforcing existing inequalities in the administration of justice. Errors in AI-driven risk assessments or predictive profiling can have serious, even irreversible consequences for individuals' liberty, raising urgent concerns about the reliability and safety of such technologies when used without adequate oversight.

Closely linked to these risks are issues of procedural fairness and accountability. Unlike traditional criminal justice practices, where responsibility can be clearly attributed to individual actors, AI introduces ambiguity in determining liability when adverse outcomes occur. Questions arise as to whether responsibility lies with the developers of the technology, the law enforcement officers using it, or the institutions that deploy it. This lack of clarity can hinder effective redressal mechanisms and weaken public confidence in the justice system.

The principles of trust, transparency, and explainability become essential components of AI governance in this context. Accused persons, defence counsel, and courts must be able to understand how AI systems arrive at particular assessments, especially in critical areas such as bail determinations, risk profiling, and evidence evaluation. The "black box" nature of many AI

models presents a significant challenge to due process and the right to a fair trial. Without such safeguards, the adoption of AI in criminal justice risks undermining the foundational values of the legal system it is meant to serve.

Effective governance mechanisms must also promote equitable application by ensuring that AI technologies are not disproportionately deployed against marginalized communities. A comprehensive policy framework should therefore be adaptive, establishing clear standards for safety, efficacy, transparency, and ethical use, while remaining responsive to technological change.

IV. Existing Legal and Policy Framework in India

India's criminal justice regulatory framework is a complex interplay of statutory laws, procedural codes, and constitutional guarantees. The Bharatiya Nyaya Sanhita (BNS) 2023[1], which replaced the Indian Penal Code[2], and the Bharatiya Sakshya Adhiniyam (BSA) 2023, which replaced the Indian Evidence Act, represent the most significant recent legislative developments. The BSA defines "document" to include "electronic and digital records," with illustrations expressly treating emails, server logs, and digital-device records as documents, thereby providing a foundational evidentiary framework for algorithmic accountability.

The Information Technology Act, 2000, serves as the primary legislation governing digital transactions and cybersecurity in India. While it does not specifically address AI in criminal justice, it provides a foundational framework for data protection and electronic governance. However, these provisions are often criticized for being outdated and insufficient to address the complexities of modern AI-driven systems, particularly with respect to algorithmic accountability, informed consent, and the contestability of automated decisions.

India's Digital Personal Data Protection Act (DPDP), 2023[3], is moving toward a risk-based governance model for AI, designating certain entities as "Significant Data Fiduciaries" based on their data processing activities. These organizations are required to conduct Data Protection Impact Assessments (DPIAs) and submit to audits of their algorithmic practices. However, the DPDP Act contains provisions that allow for exemptions in cases related to law enforcement and criminal investigations, meaning privacy arguments may not always compel transparency in AI-driven policing.

India's constitutional jurisprudence provides a strong normative basis for demanding guardrails around AI. Maneka Gandhi v. Union of India[4] is foundational for reading "procedure established

[1]Bharatiya Nyaya Sanhita, No. 45 of 2023 (India), effective July 1, 2024

[2]Indian Penal Code, No. 45 of 1860 (India), repealed by Bharatiya Nyaya Sanhita, 2023.

[3]Digital Personal Data Protection Act, No. 22 of 2023 (India).

[4]Maneka Gandhi v. Union of India, (1978) 1 SCC 248 (India).

by law" in Article 21 as requiring fairness and reasonableness. The privacy verdict in Justice K.S. Puttaswamy (Retd.) v. Union of India[5] recognized privacy as a fundamental right and established a three-part test—legality, legitimate aim, and proportionality—that maps directly onto the regulatory challenge posed by AI-driven law enforcement.

Despite these developments, several gaps and challenges persist in India's regulatory framework for AI in criminal justice. The most significant is the absence of a dedicated legal framework addressing the use of AI technologies in policing, prosecution, and adjudication. There is ambiguity regarding liability in cases of AI-induced errors, and the absence of standardized protocols for validation, certification, and auditing of AI systems creates risks related to accuracy, bias, and violations of due process. Regulatory fragmentation further complicates the landscape, as multiple authorities operate with overlapping jurisdictions and limited coordination.

V. International Policy and Regulatory Approaches

Internationally, the governance of artificial intelligence in criminal justice reflects a growing emphasis on ethical, safe, and accountable deployment. The European Union has adopted one of the most comprehensive regulatory models through the General Data Protection Regulation and the Artificial Intelligence Act. The EU AI Act[6] prohibits certain uses outright, including AI systems used to assess or predict the risk of a person committing a criminal offence when based solely on profiling or personality traits. It classifies law enforcement AI as high-risk, requiring conformity assessments, human oversight, and registration in a public database.

The comparative lesson from the EU is institutional: the framework succeeded not because of any single provision, but because it integrated accountability requirements into the lifecycle of AI systems, from design through deployment through post-market monitoring. This dual framework balances innovation with fundamental rights and individual liberty, providing a model that India can adapt rather than wholesale adopt.

In the United States, constitutional protections under the Fourth and Fourteenth Amendments have been extended, through judicial interpretation, to constrain the use of certain AI-driven surveillance and profiling tools. Several U.S. jurisdictions have enacted specific legislation limiting facial recognition use by law enforcement, reflecting growing recognition that the risks of such technologies demand targeted regulatory responses.

A comparative analysis reveals common global best practices, including risk-based regulation, mandatory transparency, human oversight requirements, and lifecycle governance. For India,

[5]Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1 (India).

[6]Regulation (EU) 2024/1689 of 13 June 2024 (Artificial Intelligence Act), art. 5(1)(d).

these approaches highlight the need for a harmonized framework that integrates constitutional protections, sector-specific AI regulation, and institutional oversight, while ensuring accessibility, affordability, and ethical deployment in a diverse criminal justice ecosystem. India need not import EU text; the legal and cultural contexts differ substantially. But India can adopt the method: identify unacceptable risk categories early; regulate high-risk systems with documentation, oversight, and redress; and phase capacity-building rather than leaving systems unregulated until a crisis forces reform.

VI. Ethical and Legal Challenges in AI Criminal Justice Governance

The integration of artificial intelligence into criminal justice governance presents a complex web of ethical and legal challenges that demand careful scrutiny. One of the foremost concerns relates to algorithmic bias and discrimination. AI models are trained on historical datasets which may inherently reflect existing social and systemic biases. When such biased data is used, AI systems can produce discriminatory outcomes, particularly affecting marginalized and underrepresented groups, resulting in wrongful targeting, unequal treatment, or exclusion from beneficial interventions.

Predictive tools amplify what scholars term "cascading bias"[7]: discretionary upstream decisions: patrol deployment, stop-and-search practices, FIR registration, shape the datasets on which predictive instruments are trained, generating feedback loops across sequential decision nodes from arrest through sentencing.[8] Facial recognition exemplifies this dynamic acutely; in India, concerns regarding operational thresholds and false-match rates acquire heightened salience when algorithmic outputs are treated as sufficient grounds for reasonable suspicion.

Informed consent and procedural fairness further complicate the ethical landscape. Traditional models of due process are based on the principle that affected individuals must be able to understand and challenge the basis of decisions affecting their liberty. However, AI systems often function as "black boxes," making it difficult for practitioners, defence counsel, and courts to fully understand how specific conclusions or recommendations are derived. This directly challenges the validity of procedural safeguards enshrined in India's constitutional framework.

Liability and accountability represent significant legal dilemmas. In cases where AI systems contribute to wrongful arrests, flawed prosecutions, or unjust sentences, it becomes difficult to determine who should be held responsible, whether liability should rest with law enforcement officers who rely on AI outputs, developers who design the algorithms, or institutions that deploy these technologies. Existing legal doctrines, largely built around human agency and mens rea, are ill-equipped to address the distributed nature of responsibility in AI-assisted decision-making.

[7]Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy 8-10 (2016).

[8]Sandra G. Mayson, Bias In, Bias Out, 128 Yale L.J. 2218, 2228-30 (2019).

Cybersecurity risks pose a growing threat to AI-enabled criminal justice systems. As justice infrastructure becomes increasingly digitized, it becomes more vulnerable to cyberattacks, including data breaches, evidence tampering, and system manipulation. Compromised systems can disrupt critical justice processes, endanger civil liberties, and result in the loss or alteration of vital evidentiary data. The ethical and legal challenges of AI in criminal justice necessitate a balanced and proactive governance approach that prioritizes individual rights, fairness, accountability, and system resilience.

VII. Role of Stakeholders in AI Governance

The governance of Artificial Intelligence in the criminal justice sector is a complex and evolving process that necessitates the active participation of multiple stakeholders, each contributing distinct roles and responsibilities to ensure ethical, safe, and effective implementation. Governments and regulatory authorities occupy a central position in this ecosystem by formulating policies, enacting legislation, and establishing regulatory frameworks that guide the development and deployment of AI technologies. Their role extends beyond mere rule-making to include oversight, monitoring compliance, and addressing emerging challenges such as algorithmic bias and cross-border data flows.

Law enforcement agencies and judicial institutions serve as the primary deployers of AI technologies in the criminal justice context. Police forces, prosecution services, and courts are responsible for integrating AI tools into their workflows, ensuring that these technologies are used as supportive instruments rather than replacements for human judgment. They must also ensure that data is collected, stored, and processed in compliance with legal and constitutional standards, thereby maintaining trust in the justice system.

Technology developers and the private sector are at the forefront of innovation in AI for criminal justice, designing and deploying systems that have the potential to transform policing, prosecution, and adjudication. Their responsibilities include ensuring that AI systems are developed in a transparent, accountable, and ethically sound manner—adopting principles such as fairness, explainability, and non-discrimination in algorithm design, as well as conducting rigorous testing and validation before deployment.

Accused persons, civil society organizations, and defence counsel are equally important stakeholders in AI governance, as they represent the individuals most directly affected by AI-driven criminal justice decisions. Their involvement is essential in ensuring that AI systems respect due process, informed consent, and human dignity. Civil society groups play a crucial role in raising awareness about the implications of AI in criminal justice, advocating for individual rights, and holding both public and private entities accountable.

The judiciary plays a significant role in shaping AI governance by interpreting laws, resolving disputes, and setting legal precedents. Courts act as guardians of fundamental rights, ensuring that the deployment of AI does not infringe upon rights such as privacy, equality, and liberty.

Judicial interventions can address gaps in existing legal frameworks, provide clarity on issues of liability and accountability, and guide policymakers in refining regulations. Public-private partnerships have also emerged as a vital mechanism for advancing AI governance, enabling collaboration between governments, justice institutions, and technology companies in developing standards and building capacity.

VIII. Key Prospects of AI in Criminal Justice

Artificial Intelligence holds significant potential to enhance the efficiency, accuracy, and accessibility of the criminal justice system. One of the most notable prospects lies in improved case management and judicial efficiency. AI-powered tools can automate routine administrative tasks, process large volumes of legal documents, and assist courts in scheduling, thereby reducing backlogs and improving the speed of justice delivery. This is particularly significant in India, where the judiciary faces a chronic shortage of judicial officers relative to the volume of pending cases.

AI also offers substantial potential in evidence analysis and forensic investigation. Machine learning algorithms can process and analyze large volumes of digital evidence, including communications data, financial records, and surveillance footage, far more rapidly and systematically than human investigators. This can enhance investigative accuracy, reduce the risk of evidence being overlooked, and support more robust prosecution of complex crimes.

Predictive analytics, when properly regulated and transparently deployed, can assist law enforcement in allocating resources more effectively, identifying high-risk locations or time periods for criminal activity and enabling pre-emptive preventive measures. Similarly, AI-driven risk assessment tools, if developed with appropriate safeguards against bias, could support more consistent and evidence-based decision-making in bail, sentencing, and parole proceedings.

AI also has significant potential in legal research and access to justice. Natural language processing tools can make legal information more accessible to individuals who cannot afford legal representation, assisting them in understanding their rights and navigating justice processes. AI-powered translation and transcription services can improve access for linguistic minorities and persons with disabilities.

The integration of AI with digital evidence management systems and the Internet of Things is reshaping the evidentiary landscape of criminal proceedings. Body cameras, smart surveillance systems, and connected devices generate vast volumes of data that can be processed and analyzed using AI to support or contest criminal allegations. Properly governed, these developments can contribute to more accurate and just outcomes; without adequate oversight, they risk enabling surveillance overreach and undermining civil liberties.

IX. Challenges in Implementation of AI Policies

The implementation of artificial intelligence policies in the criminal justice sector faces several structural, technological, and socio-legal challenges that hinder their effective realization. One of the primary concerns is the absence of a uniform and comprehensive regulatory framework. In India, AI governance in criminal justice remains fragmented, with overlapping guidelines issued by different authorities and a lack of binding legislation specifically addressing AI in policing, prosecution, and adjudication.

Another significant challenge is the disparity in infrastructure and the prevailing digital divide, particularly between urban and rural areas. While advanced law enforcement agencies in metropolitan regions may have the capacity to integrate AI technologies, rural police forces and district courts often lack the necessary digital infrastructure, trained personnel, and financial resources. This imbalance risks creating a two-tier justice system in which the benefits and risks of AI are distributed unequally.

Data-related issues also pose a critical barrier. AI systems rely heavily on large volumes of high-quality, standardized, and diverse data for training and effective functioning. However, in many criminal justice systems, data is often fragmented, inconsistently recorded, or reflects historical patterns of discriminatory enforcement. Poor data quality can lead to biased or erroneous outcomes, with severe consequences for individuals' liberty and for the integrity of the justice process.

Resistance from legal professionals and law enforcement personnel further complicates implementation. Many practitioners remain skeptical about the reliability and ethical implications of AI technologies, raising concerns about professional judgment being overridden by automated systems. This resistance is often compounded by a lack of adequate training and awareness, highlighting the urgent need for capacity-building initiatives tailored to AI integration in legal and policing contexts.

Ethical ambiguity and the absence of enforceable standards remain persistent challenges. While several ethical guidelines have been proposed, they often lack legal backing and uniform enforcement mechanisms. Issues such as accountability, transparency, and fairness in AI-driven criminal justice continue to be debated without clear resolution, making it difficult to establish binding norms for responsible AI use. Addressing these challenges is essential for building a robust and trustworthy AI governance framework that ensures both innovation and the protection of fundamental rights.

X. Case Studies and Practical Insights

The application of AI in criminal justice has been demonstrated through a range of real-world deployments that illuminate both the promise and the risks of these technologies. Predictive policing tools have been deployed across several Indian states, with systems used to forecast criminal activity hotspots and allocate police resources accordingly. While proponents claim efficiency gains, civil liberties organizations have raised concerns about the opacity of these

systems, the quality of the underlying data, and the disproportionate targeting of already over-policed communities.

Facial recognition technology has been deployed at significant scale in India, including at public gatherings, railway stations, and border crossings. The National Automated Facial Recognition System (NAFRS) represents one of the most ambitious deployments of facial recognition in a law enforcement context globally. However, studies of comparable systems internationally have demonstrated significant false-match rates, particularly for individuals from minority communities, raising serious concerns about wrongful identification and arbitrary detention.

Beyond India, international case studies provide instructive precedents. In the United States, the use of risk assessment instruments such as COMPAS in bail and sentencing proceedings has been extensively litigated and debated. The landmark case of State v. Loomis raised fundamental questions about the right of defendants to contest algorithmic risk assessments used against them, with the Wisconsin Supreme Court holding that while COMPAS could be considered alongside other factors, it could not be the determinative basis for sentencing. This case illustrates the procedural challenges that AI-driven decision support tools pose for established criminal justice principles.

From a legal and policy perspective, judicial and regulatory interventions in India remain limited but are beginning to develop. Courts have addressed the admissibility of digital evidence and have invoked constitutional principles of proportionality and fairness in technology-adjacent contexts. These developments underscore the urgent need for clearer regulatory frameworks and judicial guidelines to address accountability and ensure safe deployment of AI in criminal justice, building on India's existing constitutional and legislative architecture.

XI. From Aspiration to Enforceable Guardrails: A Four-Pillar Framework

As AI systems continue to shape the future of policing, investigation, and justice, it is essential that India's legal framework evolves to keep pace, protecting citizens from the potential harms of unchecked technological power. A workable accountability framework for criminal justice rests on four pillars.

Transparency by Default for Liberty-Affecting Systems

If an AI tool materially influences detention, bail, targeting, or surveillance, its provenance, assumptions, error rates, and field performance should be disclosable, subject to narrow, reviewable exceptions. Evidence law already recognizes server logs and digital records as documents; the State should not be able to simultaneously rely on algorithmic outputs and refuse algorithmic scrutiny. Transparency obligations of this kind should be enshrined in statute, not left to judicial inference.

Mandatory Auditability and Logging

Mandatory logs, version control, and retention rules convert "accountability" from rhetoric into proof. This aligns with the DPDP Rules' move toward due diligence for "algorithmic software" and annual DPIAs and audits for Significant Data Fiduciaries, even where policing exemptions may limit direct applicability. The standard should be clear: if the State relies on a system to deprive or constrain liberty, the State must be capable of explaining how that system works.

Contestability and Procedural Fairness

Due process for automated predictions is a well-developed legal concept internationally.[9] Individuals exposed to adverse algorithmic scoring should have meaningful opportunities to challenge decisions and demand transparency about the basis of the score. In India, Maneka Gandhi and Puttaswamy provide constitutional grounding for insisting that liberty-affecting procedures must be fair, reasoned, and proportionate. These doctrines should be operationalized through statutory rights to explanation and challenge, supported by an accessible administrative mechanism.

Institutional Accountability for Procurement and Use

Accountability must attach not only to developers but to the institutions that procure and deploy algorithmic systems. Procurement rules should require pre-deployment evaluations, independent testing, vendor disclosure commitments, and post-deployment monitoring; especially for systems that risk reinforcing discrimination. A licensing or certification regime for high-risk AI tools used in law enforcement, modeled on the DPDP Act's Significant Data Fiduciary framework, would provide a workable institutional mechanism.

XII. Future of AI Governance in Criminal Justice

The future of AI governance in criminal justice is expected to evolve through a dynamic interplay of technological advancement, regulatory innovation, and ethical standard-setting. One of the most significant emerging trends is the shift from principle-based guidelines to enforceable, risk-based regulatory frameworks. Governments and international bodies are increasingly adopting sector-specific regulations that distinguish criminal justice AI from general-purpose AI due to its direct impact on individual liberty. This includes mandatory requirements for transparency, explainability, and human oversight, particularly in policing, bail, sentencing, and evidence evaluation.

Another key trend is the growing importance of data governance and sovereignty. As AI systems rely heavily on large-scale criminal justice data, countries are moving towards stricter data protection regimes that balance law enforcement needs with individual rights. Future governance models are likely to integrate cybersecurity standards, interoperable data frameworks, and real-time monitoring systems to prevent misuse and ensure compliance. There is also a visible

[9]Danielle Keats Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249, 1253 (2008).

movement towards "responsible AI," where trust, fairness, and safety are embedded into system design rather than treated as afterthoughts.

The role of interdisciplinary approaches will become increasingly central in shaping AI governance. Effective regulation of criminal justice AI cannot rely solely on legal expertise; it requires collaboration between technologists, legal professionals, criminologists, ethicists, and policymakers. Such collaborative governance models will help bridge the gap between innovation and regulation, facilitating safer and more effective deployment in justice systems.

In the Indian context, the future of AI governance in criminal justice appears both promising and challenging. India is moving towards a structured governance framework that balances innovation with constitutional safeguards, as reflected in its emerging AI governance guidelines. Anticipated legal developments include comprehensive data protection legislation, sector-specific AI regulations, and clearer liability frameworks addressing the roles of developers, law enforcement agencies, and judicial institutions. India is likely to adopt a hybrid regulatory model that combines governmental oversight with industry self-regulation and technical standards, while remaining anchored in its constitutional commitment to fairness, equality, and the rule of law.

XIII. Conclusion

India's criminal law can meet the algorithmic age without abandoning foundational principles. The constitutional guarantees of Articles 14, 19, and 21, interpreted through the prism of Maneka Gandhi and Puttaswamy, provide normative anchors. The BSA's recognition of digital records as documentary evidence provides a procedural foundation. The DPDP Act's risk-based framework provides an institutional template. What is missing is the integrative legislative will to bind these elements together into a coherent regime of algorithmic accountability.

The urgency is not hypothetical. Predictive policing tools are already in use across several Indian states. Facial recognition systems have been deployed at public gatherings and border crossings. Automated case management systems are reshaping how courts process filings. Each of these applications carries accountability gaps that existing law does not adequately close.

This study has demonstrated that while AI holds immense potential to enhance efficiency, accuracy, and accessibility in the criminal justice system, its deployment without a robust and coherent policy framework poses significant risks to fundamental rights and the integrity of justice processes. Key findings reveal that issues such as algorithmic bias, lack of transparency, unclear liability frameworks, and inadequate contestability mechanisms continue to challenge the safe and equitable use of AI technologies. A critical evaluation of existing legal frameworks indicates that current regulatory mechanisms are not fully equipped to handle the complexities of AI systems operating within the criminal justice context.

In light of these challenges, there is a pressing need for a balanced, adaptive, and ethical governance model that can effectively regulate AI in criminal justice while fostering innovation.

Such a model must be grounded in principles of accountability, transparency, and inclusivity, ensuring that technological advancements do not compromise individual rights or deepen existing inequalities. Algorithmic accountability must be treated as a priority to ensure that AI enhances, rather than undermines, justice in India. This is not merely a legal imperative: it is a social one. The communities most exposed to AI-driven policing are those with the fewest resources to challenge its errors. A legal framework that tolerates opacity in the name of efficiency imposes its costs most heavily on those least able to bear them. In constitutional democracies, machines do not bear responsibility, institutions do. India's institutions must now demonstrate that they are equal to the task.

References

1. Bharatiya Nyaya Sanhita, No. 45 of 2023 (India), effective July 1, 2024.

2. Bharatiya Sakshya Adhiniyam, No. 47 of 2023 (India).

3. Digital Personal Data Protection Act, No. 22 of 2023 (India).

4. Information Technology Act, No. 21 of 2000 (India).

5. Maneka Gandhi v. Union of India, (1978) 1 SCC 248 (India).

6. Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1 (India).

7. Regulation (EU) 2024/1689 of 13 June 2024 (Artificial Intelligence Act).

8. Council Regulation (EU) 2016/679 (General Data Protection Regulation).

9. NITI Aayog, National Strategy for Artificial Intelligence #AIforAll (2018).

10. NITI Aayog, Responsible AI for All: Approach Document for India (2021).

11. World Health Organization, Ethics and Governance of Artificial Intelligence for Health (2021).

12. O'Neil, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (2016).

13. Mayson, Sandra G. "Bias In, Bias Out." 128 Yale Law Journal 2218 (2019).

14. Citron, Danielle Keats. "Technological Due Process." 85 Washington University Law Review 1249 (2008).

15. Mittelstadt, B.D., et al. "The Ethics of Algorithms: Mapping the Debate." Big Data & Society (2016).