Bharati Vidyapeeth
(Deemed to be University), Pune, India
New Law College, Pune
Re-accredited with 'A++' Grade by 'NAAC'
'Category - I' Status of University by UGC
Ranked 59th by NIRF

Founder Chancellor
Dr. Patangrao Kadam

BHARATI LAW REVIEW

Volume
Special Issue
June- August 2025

BHARATI LAW REVIEW

ISSN : 2278-6996
(In Collaboration With Law Artigence Cell)

THEME
ARTIFICAL INTELLIGENCE RESHAPING SOCIETY & LAW

LAW IS THE KING OF KINGS

SPECIAL ISSUE

Bharati Law Review

Bharati Law Review

(B.L.R.)
Quarterly Journal

Volume
Special Issue
June- August 2025
Free Distribution

BHARATI VIDYAPEETH (TO BE DEEMED UNIVERSITY)

NEW LAW COLLEGE, PUNE
Estd. 1978

Reaccredited with 'A^{++}' Grade by NAAC

Educational Complex, Erandwane, Paud Road,
Pune - 411 038, Maharashtra, India
Tel.: (020) 25444616 Fax: (020) 25455854
www.bvpnlcpune.org

Disclaimer

The views expressed by the various contributors of the articles are their own opinions and do not represent the Institute's opinion. No part of this publication may be reproduced or transmitted in any material form (including storing it in electronic medium, photo copy, or any other means which culminated in unauthorized duplication of the contents) without permission of the copyright holder viz., Bharati Vidyapeeth Deemed University, New Law College, Pune.

Printed at
Bharati Printing Press, BVDU, More Vidyalaya, Erandwane, Paud Road, Pune - 411 038.

Editor-in-Chief

"Artificial intelligence would be the ultimate version of Google. The ultimate search engine that would understand everything on the web. It would understand exactly what you wanted and it would give you the right thing. We're nowhere near doing that now. However, we can get incrementally closer to that and that is basically what we work on."

—Larry Page

Artificial Intelligence (AI) is revolutionizing a lot of things in society as well as impacting the legal profession resulting in some quite significant alterations to our social systems, daily lives and the practice of law. With A.I. transforming different areas of society viz., it is streamlining business processes, introducing innovating new technology and increasing efficiency in areas such as finance, transportation, retail and even legal industry remaining untouched, this change is bringing about a demand for new legislation to cope up with the changes.

As A.I. has started taking a role in making decisions, legal systems are trying to keep up with new regulations, protect individual rights, tackle A.I.-driven discrimination and in maintaining transparency in the legal process. A.I.-powered chatbots and virtual assistants are changing the game for law firms by handling client interactions, managing inquiries, scheduling appointments and in providing 24/7 support.

In the case of contracts and information, A.I. digests a vast amount of data rapidly, identifying compliance and risk matters with incredible precision. This actually maximizes the process of due diligence for legal professionals. A.I. tools render legal research, precedent analysis and review of case laws considerably more effective, reducing the amount of time attorneys have to devote to mundane tasks. Furthermore, these A.I. conclusions are capable of predicting the results of litigation which is crucial for strategy formulation.

There also exist A.I. platforms that automate contract drafting and legal documents, eliminating the potential for human error. This makes it possible for lawyers to focus more on the difficult sides of judgment and negotiation. In addition to that, A.I.-powered Online Dispute Resolution (ODR) platforms are facilitating easy access to justice for individuals. They provide faster, lower-cost substitutes for the conventional court system, assisting in decongesting courts and reaching out to disadvantaged segments of society.

All this A.I. processing is enhancing the way law firms manage contracts and detect compliance risks. Although A.I. encourages greater transparency and predictability in judicial rulings and can potentially reduce bias there remain concerns about accountability and fairness. There are still questions about how past biases in training data can influence A.I. results.

Therefore, this special issue of Bharati Law Journal on Artificial Intelligence transforming society and law is advancing the argument that A.I. is undoubtedly transforming society, streamlining and enhancing daily processes and posing significant ethical and regulatory issues. It's clear that A.I. not merely transforming the way we live but also the way we govern, settle disputes and provide justice—entrancing a new era with digital data for society as well as the law. Our legal system and society are becoming more governed by A.I. with professionals, technologists and regulatory authorities continuously having to conform to the new realities.

Dr. Ujwala Bendale,
Dean & Principal I/c,
BVDU, NLC, Pune.

Editorial

"Of Wheels, Wires, and Will: Can Artificial Intelligence Tame the Infinite Desires of Man?"

Since the dawn of time, nature fulfilled every need and desire of humankind with boundless generosity. When scarcity crept in, humanity responded with invention—tools, machines, and mechanisms—to keep pace with its ever-growing aspirations. Among these, the wheel marked a revolutionary moment in our evolution, propelling civilization forward at a speed previously unimagined across millennia of slow progress.

From that moment, human desire multiplied, and so did the machines to fulfill it. Now, in the 21st century, we stand before what many hail as the next great leap—Artificial Intelligence. Is AI the most transformative and illuminating innovation of our era? If so, a deeper question arises: Can AI truly keep up with the ever-renewing hunger of human desire? Or will it, instead, breed intellectual dependency and cognitive stagnation? Emerging research hints at a troubling trend: frequent users of AI tools like Chat GPT may experience a decline in their own creative cognition, as if the brain, relieved of its duties, begins to fall silent. The relationship between AI and Law today resembles an uneasy alliance—like that of a crocodile and a sparrow, mutually benefiting yet fundamentally alien. Yet human desire is not always gentle; it can be restless, persistent, and difficult to restrain, untamable, insatiable, and resistant to control. So the final question lingers like a storm on the horizon:

Can Artificial Intelligence cage the chaos of human craving—or will it, like all before it, be shaped and consumed by it?

Dr. Jaykumar Bhongale
Associate Professor & Editor of BLR

CONTENT

Bharati Law Review
June- August 2025

ARTIFICIAL INTELLIGENCE AND CONSTITUTIONAL EQUILIBRIUM: BALANCING INNOVATION WITH INDIVIDUAL RIGHTS

Dr. Ujwala Bendale [1]

Abstract:

The rise of artificial intelligence (AI) has brought about a significant shift in constitutional governance globally, urging societies to strike a balance between innovation and protection of fundamental rights. This paper examines the intersection of AI technologies with constitutional principles, focusing on privacy, due process, equality, and freedom of expression. Drawing on diverse legal systems, it explores landmark cases and legislative developments—such as the European Union's GDPR—as responses to AI's impact on individual rights.

Drawing upon a diverse array of constitutional contexts, this research investigates how AI applications intersect with core constitutional principles, such as privacy, equality, due process, and freedom of expression. By examining landmark cases and legislative responses, it illuminates the evolving jurisprudence surrounding AI's impact on constitutional rights. For instance, the European Union's General Data Protection Regulation (GDPR) stands as a pioneering effort to safeguard individual privacy amidst the proliferation of AI-driven data processing.[2] Moreover, this paper delves into the complex dynamics of AI governance, scrutinizing the roles of governmental institutions, regulatory bodies, and international cooperation

[1] Dean and Principal, New Law College, Bharati Vidyapeeth (Deemed to be) University, Pune.
[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data

mechanisms in shaping AI policies[3]. It also probes the challenges of accountability and transparency in AI decision-making processes, emphasizing the need for robust legal frameworks to mitigate potential biases and ensure algorithmic fairness[4]. Furthermore, the research explores emerging legal frontiers in the realm of AI, such as the notion of "digital personhood" and the implications for constitutional rights[5].

Keywords:
Artificial Intelligence (AI), Constitutional Dynamics, Governance, Rights, Legal Frameworks, Privacy, Accountability, Transparency, Algorithmic Fairness, Digital Personhood, Deepfake Technology, Ethical Considerations.

Introduction:

The age of artificial intelligence (AI) presents a profound juncture in human history, where technological advancements intersect with constitutional principles, reshaping the fabric of governance, rights, and legal frameworks[6]. As AI permeates various aspects of society, from healthcare and finance to criminal justice and entertainment, its transformative potential brings both promises and perils. ‘

At the core of this inquiry lies the tension between technological innovation and the protection of individual rights. The emergence of AI technologies, powered by machine learning algorithms and vast datasets, challenges traditional legal paradigms, raising fundamental questions about privacy,

[3] Floridi, L. (2019). Soft Ethics, the Governance of the Digital and the General Data Protection Regulation. Philosophy & Technology, 32(1), 1–7

[4] Diakopoulos, N. (2016). Accountable Algorithmic Decision-Making: Considerations for a Normative Framework. Proceedings of the Conference on Fairness, Accountability, and Transparency, 278–292.

[5] Calo, R. (2017). Artificial Intelligence Policy: A Primer and Roadmap. SSRN Electronic Journal.

[6] Waldrop, M. M. (2018). "A Brief History of Artificial Intelligence." Nature, 557(7705), 55-60

equality, due process, and freedom of expression. These questions are not merely theoretical but have tangible implications for individuals, communities, and societies at large.

Furthermore, the integration of AI blurs traditional legal boundaries, giving rise to novel concepts such as "digital personhood" and reshaping our understanding of legal rights in the digital age. Issues surrounding AI-generated content, such as deepfake technology, underscore the urgency of adaptive legal frameworks capable of preserving democratic principles and human dignity.

By analyzing landmark cases, legislative responses, scholarly debates, and emerging trends, it seeks to provide insights into the evolving landscape of AI governance, rights, and legal frontiers. Ultimately, this research underscores the imperative of balancing technological progress with the protection of fundamental rights in an AI-driven world.

Research Methodology:

This study adopts a qualitative approach to explore the constitutional dynamics in the era of artificial intelligence (AI). The methodology comprises three primary components: case studies, legal analysis, Ethical Considerations along with other aspects.

1. **Case Studies:**

The research conducts in-depth case studies to examine specific instances where AI intersects with constitutional principles in various domains such as privacy, freedom of expression, and due process. These case studies provide contextualized insights into the practical implications of AI technologies on governance and rights.

2. **Legal Analysis:**

A critical aspect of the methodology involves a comprehensive legal analysis of statutes, regulations, and judicial decisions related to AI and constitutional rights. This analysis aims to identify legal precedents, interpret statutory provisions, and assess the application of constitutional principles in AI-related cases, thereby elucidating the evolving legal landscape surrounding AI governance and rights.

3. **Ethical Considerations:**

Throughout the research process, ethical considerations are carefully observed, particularly concerning data privacy, informed consent, and the responsible conduct of research involving human subjects. The study adheres to ethical guidelines and principles of research integrity to ensure the ethical treatment of participants and the responsible dissemination of findings.

The rapid advancement of artificial intelligence (AI) presents pressing challenges to existing constitutional frameworks, particularly concerning fundamental rights and governance. However, current research lacks a comprehensive understanding of AI's specific implications on constitutional dynamics and the necessary mechanisms to address emerging challenges effectively. This gap hampers policymakers, legal practitioners, and stakeholders in navigating the complex terrain of AI governance and rights protection, necessitating focused research to bridge this knowledge divide.

Hypothesis:

This study hypothesizes that the integration of artificial intelligence (AI) into society poses significant challenges to existing constitutional frameworks, particularly in terms of protecting fundamental rights and ensuring effective governance. Furthermore, it is posited that the rapid advancement of AI technology outpaces the development of appropriate legal and regulatory frameworks, leading to gaps in rights protection and accountability mechanisms. Lastly, the hypothesis suggests that a comprehensive understanding of the implications of AI on constitutional dynamics is essential for informing policy and legal responses to ensure the effective protection of rights in the AI era.

Jurisprudence Aspect

The jurisprudence of the interplay between artificial intelligence (AI) and constitutional principles encompasses several critical dimensions. Firstly, it scrutinizes the protection of fundamental rights guaranteed by constitutions, including privacy, freedom of expression, and due process, in the context of AI technologies. Legal precedents and decisions offer insights into balancing AI innovation's benefits with safeguarding individuals' rights and liberties. For instance, recent cases involving AI-powered surveillance or algorithmic decision-making have prompted courts to reassess the scope of privacy rights and the limits of government intrusion.

Secondly, jurisprudence addresses algorithmic accountability concerns, evaluating the legal responsibility of AI developers, users, and decision-makers for potential bias, discrimination, and unfair outcomes arising from algorithmic decisions[7].

Additionally, the governance and regulatory frameworks surrounding AI come under jurisprudential scrutiny[8]. Courts assess the constitutionality of legislative measures aimed at regulating AI, such as data protection laws and algorithmic transparency requirements, using legal principles of separation of powers and procedural due process to guide their analysis. They may also

[7] Jones, R. (Year). "Algorithmic Accountability in the Age of Artificial Intelligence." Journal of Law and Technology, 10(2), 45-67.

[8] Brown, S. (Year). "Regulatory Challenges in the Governance of Artificial Intelligence." Regulatory Review, 20(4), 256-275.

review administrative decisions related to AI governance to ensure compliance with constitutional norms and principles.

Furthermore, jurisprudence plays a pivotal role in interpreting constitutional provisions in light of emerging AI challenges and technological advancements[9]. Courts adapt legal doctrines and precedents to address novel issues like the rights of AI-generated content and digital personhood, ensuring that legal frameworks effectively safeguard rights and adapt to the changing technological landscape[10]. Through judicial interpretation, legal analysis, and precedent-setting decisions, jurisprudence shapes the legal framework surrounding AI, informing policy and regulatory responses to ensure AI technologies' compatibility with constitutional values and principles[11]

Constitutional Aspect

The interplay between artificial intelligence (AI) and constitutional principles is not only significant but also intricate, given the diverse and dynamic nature of Indian society. Beyond the fundamental rights guaranteed by the Constitution, such as privacy, freedom of speech, and equality, constitutional analysis delves into the broader ethos and values embedded in the Indian Constitution.

One crucial aspect is the principle of social justice enshrined in the Preamble and various provisions of the Constitution. AI technologies have the potential to either exacerbate existing social inequalities or contribute to bridging the gap, depending on how they are designed, deployed, and regulated. Therefore, constitutional scrutiny involves ensuring that AI applications

[9] White, A. (Year). "Interpreting Constitutional Rights in the Age of Artificial Intelligence." Harvard Law Review, 105(4), 332-349.

[10] Green, B. (Year). "Digital Personhood: Legal and Ethical Implications." Ethics and Information Technology, 25(1), 78-95.

[11] Miller, C. (Year). "The Role of Jurisprudence in Shaping AI Governance." AI and Society, 30(2), 145-167.

promote inclusive growth, equitable opportunities, and social welfare, in line with the constitutional mandate of achieving a just society.

Moreover, the Indian Constitution mandates the protection of cultural diversity and the promotion of cultural and educational rights for minorities. In the context of AI, constitutional analysis extends to issues of cultural sensitivity, linguistic diversity, and the preservation of indigenous knowledge systems. AI technologies must respect and accommodate India's rich cultural tapestry, ensuring that they do not undermine or marginalize minority cultures or languages.

Furthermore, the constitutional aspect of AI governance intersects with principles of federalism, as India is a federal republic with a division of powers between the central and state governments. The distribution of legislative and regulatory authority over AI-related matters requires careful constitutional analysis to ensure coherence, coordination, and effective governance at both levels of government.

Additionally, constitutional scrutiny encompasses the principle of technological sovereignty, which emphasizes India's ability to shape and control its technological future in alignment with national interests and values. As AI technologies become increasingly central to India's development and security agendas, constitutional analysis guides policymakers in formulating strategies to harness AI for national development while safeguarding sovereignty and autonomy in the digital domain.

In summary, the constitutional aspect of the interplay between AI and legal frameworks in India is multifaceted, touching upon issues of social justice, cultural diversity, federalism, and technological sovereignty. By upholding constitutional values and principles, constitutional analysis guides the formulation of AI governance frameworks that foster inclusive growth, respect cultural pluralism, ensure effective governance, and uphold India's sovereignty in the digital age.

K.S Puttaswamy vs Union of India (2017)

The case of *Justice K.S. Puttaswamy (Retd.) v. Union of India*[12], commonly known as the "Aadhaar judgment," marks a significant milestone in Indian jurisprudence regarding the right to privacy. The petitioners challenged the constitutional validity of the Aadhaar project, a biometric identification system initiated by the Indian government, arguing that it infringed upon the right to privacy guaranteed under Article 21 of the Indian Constitution.

The Supreme Court, in its landmark judgment, recognized the right to privacy as a fundamental right intrinsic to the right to life and personal liberty under Article 21[13] of the Indian Constitution. The court held that privacy encompasses informational privacy, bodily integrity, and decisional autonomy, among other aspects, and is essential for the exercise of other fundamental rights.

This recognition of the right to privacy has profound implications for the regulation of artificial intelligence (AI) in India. Several aspects of AI governance intersect with constitutional principles, as highlighted below:

1. **Right to Privacy**: The Puttaswamy judgment establishes the right to privacy as a cornerstone of individual autonomy and dignity. This right is directly relevant to AI technologies, which often involve the collection, processing, and analysis of vast amounts of personal data. AI systems, particularly those

[12] K.S. Puttaswamy (Retd.) and Another v. Union of India and Others (2017) 10 SCC 1
[13] Article 21 of the Indian Constitution of India Act 1 of 1950 .

driven by machine learning algorithms, raise concerns about privacy infringement and surveillance, necessitating robust data protection measures.

2. **Data Protection**: The court's emphasis on the need for a robust data protection regime aligns with the challenges posed by AI-driven technologies. AI systems rely heavily on data, and ensuring data security and privacy is paramount to prevent unauthorized access, data breaches, and misuse. The judgment underscores the importance of regulating the collection, storage, and use of personal data to safeguard individual rights and prevent abuse by state and non-state actors.

3. **Governmental Surveillance**: The Puttaswamy judgment imposes restrictions on governmental surveillance and the use of technology for mass surveillance purposes. AI-driven surveillance technologies, such as facial recognition and predictive policing algorithms[14], raise concerns about privacy violations, discrimination, and infringement of civil liberties. The judgment underscores the need for transparency, accountability, and judicial oversight in regulating surveillance technologies to prevent abuse and protect democratic values.

4. **Autonomy and Decisional Privacy:** The recognition of decisional autonomy as a core component of privacy rights has implications for AI technologies involving automated decision-making systems. AI algorithms may impact individuals' autonomy and decision-making processes, raising concerns about bias, discrimination, and lack of accountability. Safeguards are necessary to ensure that AI systems respect individuals' autonomy and prevent undue influence or harm.

[14] Facial recognition involves identifying individuals from facial features captured in images or videos, while predictive policing algorithms use data analysis to forecast potential criminal activity.

5. **Democratic Values:** The Puttaswamy judgment reaffirms the importance of democratic values, including transparency, accountability, and the rule of law, in governing technological advancements. AI governance must uphold constitutional principles and democratic norms to ensure that technological innovations serve the public interest and protect individual rights and freedoms. The judgment emphasizes the need for ethical and responsible AI deployment that promotes democratic values and protects human rights.

R. v. Chief Constable of South Wales Police (2020)

In the case of R. v. Chief Constable of South Wales Police[15], the UK Supreme Court addressed the use of facial recognition technology (FRT) by law enforcement authorities and its compatibility with constitutional and human rights principles. The case centered on the deployment of FRT by the South Wales Police in public spaces, raising significant concerns about privacy, data protection, and the rule of law.

The introduction of FRT by law enforcement agencies represents a pivotal intersection between artificial intelligence (AI) and legal frameworks. FRT systems analyze biometric data from individuals captured in real-time through surveillance cameras, enabling the identification and tracking of individuals in public spaces. This technology has significant implications for privacy rights, freedom of movement, and the presumption of innocence, posing challenges for legal systems tasked with safeguarding fundamental rights in the digital age.

In this case, the petitioners argued that the use of FRT by the South Wales Police infringed upon their right to privacy under the UK Human Rights Act. They contended that the indiscriminate collection and analysis of biometric

[15] R v. Chief Constable of South Wales Police [2020] UKSC 25.

data without individual consent amounted to unlawful surveillance and violated their fundamental rights. The case raised questions about the legality, proportionality, and necessity of FRT deployments, particularly concerning its potential for mass surveillance and discriminatory outcomes.

The Supreme Court's judgment in this case underscored the importance of upholding constitutional and human rights principles in regulating AI technologies. While recognizing the legitimate aims of law enforcement in maintaining public safety and preventing crime, the court emphasized the need for clear legal frameworks and robust safeguards to protect individual privacy and prevent abuses of power.

The court's ruling highlighted the inherent risks and challenges associated with the use of AI technologies in law enforcement, particularly concerning bias, discrimination, and lack of transparency. It called for greater accountability, oversight, and transparency in the deployment of FRT, including clear policies on data retention, public consultation, and independent judicial review.

The case serves as a crucial precedent for future AI and law interactions, emphasizing the need for legal frameworks that uphold fundamental rights, promote transparency and accountability, and ensure that technological innovations serve the public interest while respecting human dignity and autonomy. It underscores the role of the judiciary in safeguarding democratic values and protecting individuals from the potential harms of AI-driven surveillance and law enforcement practices.

Administrative Aspect

In the administrative aspect of the intersection between artificial intelligence (AI) and the law, several key considerations come into play. One primary challenge lies in developing regulatory frameworks to govern the use of AI in legal processes. These frameworks must address complex issues such as data protection, privacy rights, algorithmic transparency, and accountability. Government agencies bear the responsibility of drafting and enforcing regulations that strike a balance between fostering innovation and safeguarding individual rights. Government agencies, such as the Federal Trade Commission[16] (FTC) in the United States or the Information Commissioner's Office (ICO) in the United Kingdom, develop regulations governing the use of AI in legal contexts. For instance, the GDPR (General Data Protection Regulation)[17] in the European Union sets strict guidelines for the processing of personal data, including data used in AI systems employed by legal entities.

Ethical guidelines play a crucial role in guiding the administrative efforts related to AI in legal contexts. Administrative bodies often establish principles such as fairness, transparency, accountability, and non-discrimination to govern the use of AI. These guidelines serve as a foundation for ensuring that AI applications uphold the rule of law and respect fundamental rights and freedoms. Government agencies work to integrate these ethical considerations into their regulatory frameworks and operational practices.

Accountability mechanisms are essential to address instances of AI bias, error, or misuse in legal proceedings. Government agencies are tasked with implementing oversight bodies, conducting audits of AI systems, and providing avenues for redress for individuals affected by algorithmic decisions. These mechanisms serve to promote transparency, accountability,

[16] The Federal Trade Commission (FTC) in the United States is a regulatory agency responsible for promoting consumer protection and competition in the marketplace.
[17] The General Data Protection Regulation (GDPR) is a comprehensive privacy and data protection law enacted by the European Union (EU) to safeguard the personal data of individuals within the EU and the European Economic Area (EEA).

and trust in AI-driven legal processes, thereby enhancing public confidence in the justice system.For example, the Legal Services Regulatory Authority (LSRA) in Ireland may conduct audits of law firms' AI systems to ensure compliance with data protection regulations and ethical standards.

Training and education initiatives are vital for building capacity among legal professionals and government officials to navigate the complexities of AI in legal contexts. Government agencies invest in programs aimed at raising awareness of AI capabilities and limitations, promoting best practices for integrating AI into legal processes, and fostering a culture of responsible AI use. By providing training and education, administrative bodies empower stakeholders to leverage AI effectively while mitigating risks and ensuring compliance with legal and ethical standards.

Research and development efforts are another key aspect of the administrative landscape surrounding AI and the law. Government agencies may allocate resources to fund interdisciplinary research projects, support collaborations between AI experts and legal scholars, and foster innovation in areas such as legal research, contract analysis, and case prediction. By investing in research and development, administrative bodies drive advancements in AI technologies for legal applications, enhancing the efficiency and effectiveness of legal processes.The National Science Foundation (NSF) in the United States[18] funds research projects exploring the intersection of AI and law. For example, NSF grants support interdisciplinary studies on AI-powered predictive analytics in criminal justice, examining issues such as fairness, bias, and due process concern

[18] The National Science Foundation (NSF) in the United States is a government agency dedicated to funding scientific research and promoting innovation across various fields of science and engineering.

International cooperation is essential for addressing the global challenges posed by AI in legal contexts. Government agencies engage with international organizations, regulatory bodies, and other governments to share best practices, harmonize standards, and address cross-border legal issues related to AI. By collaborating on regulatory initiatives and exchanging knowledge and expertise, administrative bodies strengthen the global governance framework for AI and contribute to the responsible and ethical use of AI in the legal domain.

Overall, the administrative aspect of AI and the law requires proactive measures to develop regulatory frameworks, establish ethical guidelines, implement accountability mechanisms, promote training and education, foster research and development, and facilitate international cooperation. By addressing these challenges, government agencies can harness the potential of AI to enhance legal processes while upholding the rule of law and protecting individual rights and freedoms.

Findings

1.**Constitutional Significance:** The research underscores the profound significance of constitutional principles in governing the deployment and use of artificial intelligence (AI) technologies within legal frameworks. Constitutional rights, including privacy, autonomy, and due process, serve as foundational pillars that guide the development of regulations and policies concerning AI in legal contexts.

2. **Complex Regulatory Landscape:** The findings reveal the complexity of establishing regulatory frameworks to govern AI in legal domains. Addressing issues such as data protection, algorithmic transparency, and accountability poses significant challenges for policymakers and regulatory bodies. Crafting regulations that strike a balance between fostering innovation

and safeguarding individual rights requires careful consideration of diverse and evolving legal and technological landscapes.

3. **Governmental Role and Responsibility:** Government agencies play a critical role in shaping AI governance through policy formulation, regulation enforcement, and international cooperation. These agencies are tasked with developing and implementing measures to address regulatory gaps, promote responsible AI use, and uphold constitutional principles. Their actions and decisions have far-reaching implications for ensuring that AI technologies align with legal norms and respect human rights.

4. Cases: As highlighted in landmark cases such as K.S. Puttaswamy v. Union of India and R. vs. Chief Constable of South Wales Police, it plays a crucial role in shaping AI governance within legal frameworks. These cases underscore the importance of fundamental rights, including privacy, autonomy, and due process, in regulating AI technologies and ensuring their alignment with constitutional norms.

5. Ethical Imperatives: Ethical considerations emerge as a central theme in AI governance, particularly within legal systems. The research highlights the importance of establishing ethical guidelines to promote fairness, transparency, and accountability in AI-driven legal processes. Ethical frameworks help mitigate risks associated with bias, discrimination, and misuse of AI technologies, enhancing public trust and confidence in the justice system.

Analysis

Exploring the intricate relationship between artificial intelligence (AI) and constitutional dynamics within legal frameworks demands a thorough examination of the complex implications and evolving nature of this intersection.

AI technologies fundamentally challenge traditional approaches to legal interpretation and application by introducing innovative methods of data analysis, decision-making, and automation. This shift in paradigm necessitates a reassessment of established legal principles and constitutional rights, particularly concerning privacy, autonomy, and due process.

A significant challenge arises in reconciling the tension between technological progress and constitutional adherence. AI systems, driven by algorithms and machine learning, have the capacity to process vast amounts of data and generate insights at unprecedented speeds. While offering opportunities for efficiency and innovation within legal processes, these capabilities also raise concerns regarding algorithmic bias, discrimination, and transparency.

The landmark case of K.S. Puttaswamy v. Union of India serves as a poignant example of the judiciary's role in safeguarding fundamental rights amidst technological advancements. The Supreme Court's acknowledgment of the right to privacy as a fundamental right intrinsic to the right to life and personal liberty under Article 21 of the Indian Constitution underscores the necessity to adapt constitutional principles to the digital age. This ruling establishes a precedent for future cases involving AI technologies, emphasizing the judiciary's duty to ensure that legal frameworks effectively protect individual rights in an increasingly digital society.

Similarly, the case of R. v. Chief Constable of South Wales Police sheds light on the ethical and legal implications of deploying AI technologies, such as facial recognition, in law enforcement settings. The UK Supreme Court's consideration of the right to privacy under the Human Rights Act underscores the importance of balancing public safety concerns with individual privacy rights. This case underscores the judiciary's responsibility in scrutinizing the use of AI technologies by government agencies, ensuring that such deployments align with constitutional norms and legal standards.

A critical analysis of these cases reveals the intricate interplay between technology, law, and society. While AI technologies hold promise in transforming legal processes, enhancing access to justice, and improving decision-making outcomes, they also present significant challenges in terms of accountability, transparency, and fairness.

Addressing these challenges necessitates a holistic approach that integrates legal, ethical, and technological considerations. Collaboration among policymakers, regulatory bodies, and legal professionals is essential in developing comprehensive AI governance frameworks that uphold constitutional principles while fostering innovation and serving the public interest.

Furthermore, fostering public awareness and engagement is crucial in promoting informed discourse on the implications of AI technologies for constitutional rights and legal norms. By cultivating a culture of responsible AI use and democratic deliberation, society can harness the potential of AI technologies to strengthen the rule of law and uphold democratic values in the digital age.

In essence, the intersection of AI and constitutional dynamics within legal frameworks presents both opportunities and challenges for societies worldwide. A nuanced understanding of this complex relationship is essential in navigating the evolving landscape of AI governance and ensuring that technological advancements are harnessed to uphold fundamental rights and democratic principles.

Conclusion

The convergence of artificial intelligence (AI) and constitutional dynamics within legal frameworks presents a nuanced landscape characterized by both promise and challenge for societies worldwide. This paper provides a comprehensive exploration of this intersection, drawing upon in-depth analysis of key themes and landmark cases to illuminate the transformative potential of AI in legal processes, as well as the ethical, legal, and societal implications that accompany such advancements.

Artificial intelligence technologies hold the promise of revolutionizing legal processes, offering opportunities to enhance access to justice and improve decision-making outcomes. Through the application of AI algorithms, tasks such as legal research, document analysis, and case management can be streamlined, allowing legal professionals to focus their expertise on higher-level strategic considerations. Moreover, AI-powered predictive analytics can assist in identifying patterns and trends within legal datasets, thereby facilitating more informed decision-making by judges and policymakers.

However, the transformative potential of AI in legal contexts is not without its challenges. Significant ethical considerations arise concerning the use of AI algorithms in decision-making processes, particularly in areas such as criminal justice where individual rights and liberties are at stake. Concerns about algorithmic bias, transparency, and accountability underscore the need for robust governance frameworks to ensure that AI technologies are deployed responsibly and in accordance with constitutional principles.

Central to navigating the complexities of AI governance within legal systems is the recognition and protection of fundamental rights, including privacy, autonomy, and due process. Landmark cases such as K.S. Puttaswamy v. Union of India highlight the judiciary's pivotal role in safeguarding these rights and ensuring that legal frameworks remain relevant and effective in the digital age. In this case, the Supreme Court of India affirmed the right to privacy as a fundamental right inherent in the Constitution, thereby establishing a legal precedent that has far-reaching implications for the regulation of AI technologies and data processing activities.

Similarly, cases such as R. v. Chief Constable of South Wales Police underscore the ethical and legal considerations surrounding the deployment of AI technologies in law enforcement contexts. In this case, the use of facial recognition technology by law enforcement agencies was challenged on the grounds of its potential infringement upon individual privacy rights. While the court ultimately ruled in favor of the police, emphasizing the importance of public safety concerns, the case highlights the need for a delicate balance between security imperatives and civil liberties in the development and deployment of AI technologies.

Moving forward, addressing the challenges posed by AI in legal contexts necessitates a comprehensive and multidisciplinary approach. Policymakers, regulatory bodies, legal professionals, and technologists must collaborate to develop ethical AI governance frameworks that uphold constitutional principles, foster transparency, accountability, and fairness, and promote the public interest. Such frameworks should incorporate mechanisms for ongoing monitoring and evaluation to ensure that AI technologies are deployed responsibly and in alignment with democratic values and constitutional norms.

Moreover, fostering public awareness and engagement is crucial in ensuring that AI technologies are deployed in a manner that respects individual rights and liberties. Educating the public about the potential risks and benefits of AI in legal contexts can help to build trust and confidence in the use of these technologies, thereby facilitating greater acceptance and adoption within society. Additionally, soliciting input from diverse stakeholders, including civil society organizations, academic institutions, and marginalized communities, can help to ensure that AI governance frameworks are inclusive and representative of a broad range of perspectives and interests.

In essence, the evolution of AI within legal frameworks represents a pivotal moment in the history of law and society. By embracing the opportunities presented by AI technologies while proactively addressing the associated challenges, societies can harness the transformative potential of AI to strengthen the rule of law, uphold fundamental rights, and promote justice and equality for all. However, achieving these objectives will require sustained effort and collaboration across multiple sectors, guided by a shared commitment to ethical principles and democratic values.

LEGAL SCENARIO OF ELECTRONIC CONTRACT IN THE BUSINESS WORLD: A STUDY

Dr. Rashmi Dubey[1]

Abstract*:*

The focus of this paper is **Electronic Contracts: A Creative Objective in the Business World** and it asserts that recent developments in computer, telecommunications, software, and information technology have drastically altered peoples' standards of living. Due to time and geographical limitations, contact is not any less restricted. Wider and faster than ever before, information is being delivered and received. Additionally, outmoded business models have emerged in recent years, and the owners' or shareholders' primary source of income is no longer relevant. In light of this, the researcher here specifies that the presence of e-contracts on the market satisfies the need for innovation in the traditional business segments. Existing and new businesses alike are working to develop an online identity and an e-contract stance while keeping in mind the demands of the digital age. Thus, this article supports the idea that electronic contracts—contracts that are mostly in electronic form rather than written down—are the result of a desire for convenience, speed, and effectiveness.

Index Terms: e-contract, e-signatures, Indian scenario of e-contract, International recognition of e-contract, types of e-contract, corona and e-contract[2]

Introduction-

Since contracts are now so prevalent in daily life, we frequently aren't even aware when we've signed one. Numerous aspects of our everyday life are governed by contracts, from ordering a taxi to purchasing airline tickets

[1] Assistant Professor, Bharati Vidyapeeth Deemed to be University, New Law College, Pune.

[2] Available at https://scholarticles.wordpress.com/2015/08/27/legal-recognition-to-electronic-records-in-india

online. Contract formation and performance in India are governed by the Indian Contract Act, 1872. It specifies the consequences of a contract's obligations being broken as well as how the requirements in a contract are to be carried out. Parties are able to negotiate their own terms of agreement within the framework of the Act. Limiting circumstances under which a contract may be made, carried out, and its breach enforced are covered under the Indian Contract Act. It simply offers a general description of the laws and norms that control the formation and execution of contracts. The parties to the agreement agree on the parties' rights, obligations, and other terms. In the event of a default, the court takes action to enforce the agreement. Electronic contracts were developed in response to the demands for efficiency, appropriateness, and speed. Consider a deal that an American buyer and an Indian exporter want to make. One possibility is for one party to print off two copies of the agreement, sign them, and send them by courier to the other, who will then sign both copies and send one copy back. The alternative is for the parties to meet and sign the agreement there. By merely adding both parties' digital signatures to an electronic copy of the contract, the entire deal can now be executed in a matter of seconds. In such a case, a courier delay and additional travel expenses are unnecessary. Legislators initially resisted adopting legislation to acknowledge this cutting-edge technology, but today several nations have done so.[3]

Non-Empirical Approach-

The discussion of the "Electronic Form of Contract" part is doctrinal in nature. In business and many other fields, electronic forms of contracts are increasingly becoming a necessary part of contract law. In addition to adding new provisions, the Indian Contract Act is being simplified to encourage the

[3] Available at http://jcil.lsyndicate.com/wp-content/uploads/2017/10/Rachna-Choudhary.pdf

use of such contracts. The urgent requirement is for such a contract to be accepted on a global scale.[4]

A Comprehensive Approach to E-Contracts:

Business environments can be more flexible thanks to electronic contracts in terms of place, time, space, distance, and payment. This electronic contract relates to the exchange of data, goods, and services across computer networks. It is a method of conducting business electronically, most frequently online. It is the means by which "enterprise integration" is accomplished. E-contract usage is advancing quickly along with the expansion of e-commerce. However, the execution of electronic contracts is fraught with difficulties on three levels: conceptual, logical, and practical. The newspaper industry in the United States of America is a typical illustration of such a circumstance, since many of the notable newspapers have either closed their doors or switched entirely to the internet media. Behind-the-scenes couriers and additional travel expenses are not necessary in this circumstance. Just a few basic elements must be met for the E-Contracts to be valid the submission of an offer, acknowledgement of offer, Legal Consideration and Lawful Relations.[5]

E-contract types include:

Without a face-to-face encounter between the parties, electronic contracts are agreements made through online commerce. These agreements resemble paper-based business contracts in many ways, except that electronic transactions are done and agreements are reached. The globalization of society and the development of technology have hastened the growth of e-

[4] Available at https://legalserviceindia.com/articles/ecta.htm

[5] Available at https://www.lawyersclubindia.com/articles/e-contract-and-its-validity-M-Piravi-Perumal-665.as

commerce businesses. The following list of E-Contract types is provided: the offer, consideration of legal issues, and legitimate business relationships.

Browse Wrap Agreements- By using the website, the contracting parties agree to be bound by this agreement, which is referred to as a browse wrap agreement. These are the "terms of use", "user agreement", or "terms of service", which can be accessed via links at the bottom or corner of the page. They include the user policies and terms of service of websites like Flipkart or E-bay.

Shrink Wrap Contracts-These licence agreements, by which the terms and conditions of the contract are enforced upon the contracting parties, are typically included with the software items that consumers purchase and can be found on packaging or in manuals.

Click Wrap Agreement- By selecting the "Ok" or "I agree" button, the user agrees to the end user agreement's terms and conditions, which govern the licensed use of the product. There are particular types of checks that make sure the provisions of the contract are enforceable against the parties. Following are these: It must be made clear to the parties what the user agreement or terms of service entail. It is not deemed to be an indication of the user's understanding to merely include a link to the terms on the website without calling the user's attention. The terms of the agreement should not be changed if the user has given his consent for the particular action; the changes made to the terms of the agreement must be specifically intimated to the user which provides a user to give a new consent for the modifications in the terms. Accordingly, if the user continues to use the website after receiving notice of the terms, it will be assumed that they have accepted the contract.

The user has the option to quit the website immediately if he does not accept the changes.[6]

Current status of e-contract in India:

Undoubtedly, the use of electronic contracts is growing in India. Indian Railway Catering and Tourism Corporation Limited (IRCTC) is undoubtedly India's top e-commerce site and the country's response to initiatives backed by private equity. However, there are absolutely no rules governing the payment methods that every e-commerce website must accept or the transaction fees that they may charge. It has been observed that several travel websites add extra fees to the payments made by the users. According to the websites, that is a standard procedure in the sector. As a result of concerns about anti-competitive practices, the situation becomes even more complicated. The Indian Contract of 1872 recognized conventional contracts, including oral agreements formed with the free assent of the contracting parties and for legitimate consideration with a lawful intent that are not specifically declared void. Therefore, there is no clause in this Act that forbids the enforcement of electronic agreements under the condition that they contain all of the requirements for a legitimate contract. The primary elements of a legally binding contract are thought to be the parties' free consent. E-contracts typically do not allow for negotiation. The user always has the choice of a "take it or leave it" transaction. The Indian courts have addressed the legality of e-contracts in a number of cases, including those involving contract term negotiations. According to the Supreme Court's ruling in the case of LIC India v. Consumer Education and Research Centre, "In dotted line contracts there would be no necessity for a weaker party to bargain as to assume to have equal bargaining strength. According to the conditions of the dotted line contract, he must either accept the service or leave the items. He would have

[6] Available at https://www.scribd.com/document/353290697/e-Contract

the choice of continuing to use the service indefinitely or accepting the unfair or unreasonable terms.[7]

E-contracting scenario on a global scale:

New consumer protection threats necessitate new safeguarding regulations and practices. It is important to note that improved online consumer protection would positively affect the expansion of electronic commerce and, in turn, the profitability of retailers. In general, if internet commerce is to grow, customers must be given at least the same protections as they would in the traditional market. The need for a worldwide coordinated strategy to address the problem of dispute firmness in electronic business has been highlighted by OECD Member States. Procedures for Consumer Protection in the Context of Electronic Commerce, a crucial document created by the OECD, outlines procedures for consumer protection in dispute resolution and amendments that aim to protect consumers participating in electronic commerce without creating trade obstacles. Web wrap contracts and shrink wrap contracts are web-based contracts that call for the consent of the parties and that users accept when installing software from a.c.d. ram. Also included are Digital Signature, Online Banking, and Other Online Transaction as methods of Electronic Contract. The following questions must be taken into account in order for e-contracts to be recognized: Whether an electronic contract is enforceable? Would a provider be considered to have made an offer if they posted information about their products and prices on a website? Does the legal need that agreements be reduced to signed documents apply to e-contracts? Does the interpretation, adoption, and compilation of the other existing legal standards in the context of electronic transactions apply to e-contracts?[8]

[7] Available at https://www.lawfarm.in/blogs/all-you-need-to-know-about-e--contract

[8] Available at https://www.scribd.com/document/304325104/E-contracts-Short-Project

Legal foundation for e-contracts:

E-Contracts and their legality are a key element in the development of e-commerce. The first and most important rule of business law is faithfulness and truthfulness. However, in the age of e-commerce, trust is becoming less clear since a lot of customers are exposed to the seductive marketing made by e-vendors and because some unwary customers are being duped by the false promises made by these e-vendors. E-contracts create an extremely thin line between trust and being duped because trust cannot be partial.[9]

The modern world is an "e-world" or "virtual world," propelled by e-speed, digital invention, and limitless space, where money takes the role of time. The many parties are constantly recognizing and assessing the subtleties of the legal framework around it due to the increasing importance and value of e-contracts in India and throughout the world. The e-contract business has become more complicated as a result of the involvement of numerous service providers, including a payment gateway, the primary website, the bank or card verification website, the security authorization website, and the final service provider, which may also include the shipping agent. The need for amendable it has increased as a result. There are now no clear laws or regulations in India safeguarding those who purchase and sell goods and services online. However, a number of regulations working together are attempting to control E-contract commercial transactions. The Indian Contract Act of 1872, the Consumer Protection Act of 1986, the Information Technology Act of 2000, and the Indian Copyright Act of 1957 are a few examples. E-contract business operates on the same principles as other business kinds.[10]

[9] Available at https://www.grin.com/document/42720

[10] Available at http://www.legalservicesindia.com/article/1943/A-study-of-Formation-and-challenges-of-electronic-contract-in-cyberspace.html

The attribution, acknowledgement, and dispatch of electronic records and secured electronic procedures are covered by the IT Act's provisions. The IT Act recognizes the fundamental elements of a contract, including the communication of proposals, acceptance of proposals, and, if necessary, revocation of proposals and acceptances, which may be conveyed electronically or through the use of an electronic record. The Indian Evidence Act also recognizes contracts, and it defines a "document" as any information in an electronic record that is printed on paper, stored, recorded, or reproduced on optical or magnetic media created by a computer. Such information is in accordance with the requirements of Section 65B of the Act and is admissible in any proceedings without the need for further justification or the production of the original document before the relevant authority. It is also regarded as evidence of the contents of the original or any fact stated therein for which direct testimony would be admissible.[11]

E-signatures and e-contracts:

E-signature refers to a digital file or symbol, such as a typed name or scanned pen-and-ink signature that a person adds to or inserts on a document to indicate their intent to sign it. People can electronically sign papers in a variety of methods, including as putting their name into the signature field, inserting a scanned copy of their signature, clicking a "I accept" button, or utilizing cryptographic "scrambling" technology. There are varying degrees of security among various e-signing techniques. The term "electronic signature" is typically used to refer generically to all e-signatures, including those that utilize shaky techniques. The signer's name entered into the signature field and a scan of their pen-and-ink signature are two unreliable signature types. Although these formats are still legally binding, they virtually remove any possibility of verifying that the person who signed the document was the one you wanted to sign it with. On the other hand, "digital signatures"

[11] Available at https://www.scribd.com/document/372333901/e-Contracts

are more advanced, secure electronic signatures. To authenticate the signer, they employ digital identification. The paper is then electronically sealed with the signature utilizing encryption. To create a digital signature, you don't need to be an expert in computers; a number of software tools, including DocuSign, HelloSign, Adobe, and SignNow, make the procedure quick, easy, and inexpensive.[12]

Every day, thousands of business deals are made over the Internet without any face-to-face connection between the parties. People use electronic means to carry out a variety of tasks, including buying insurance, signing real estate contracts, using credit cards, and entering into financial agreements. Despite the widespread use of electronic transactions, many individuals are still unclear as to whether e-signatures and e-contracts are legitimate, safe, and legal. E-contracts and e-signatures are, for the most part, secure and dependable ways to conduct business, which is excellent news for both businesses and customers. However, some precautions should be taken by the parties to e-contracts and e-signature agreements to guarantee the legality of their agreements.[13]

E-contract and covid 19:

Everybody's business practices have changed as a result of the global COVID-19 outbreak, including those of promoters, investment bankers, transactional lawyers, and fund managers. The signing of paperwork will be among the major obstacles in any actual transaction or business negotiation. The Indian market is adopting remote execution, however the parties are wary of this strategy and desire actual physical signatures on the documents. Physical execution, however, is virtually difficult because of continuous

[12] Available at https://www.scribd.com/document/291526593/Contracts-II-Final-Project
[13] Available at https://www.coursehero.com/file/41945010/E-CONTRACTdocx/

travel restrictions and the potential for lockdowns.[14] Adopting electronic contracts and document signing electronically is one way to address this problem. India Inc. may be propelled towards more efficient and paperless document execution methods through COVID-19.[15] Foreign signatories who do not possess digital signature or Aadhaar e-sign will not be permitted to engage in e-signing, however, as the IT Act expressly recognizes only these two forms of electronic signature. In that situation, the foreign signatory may rely on the signing method at their disposal and demonstrate its validity using evidence like email correspondence or the behaviour of the parties to indicate purpose.[16]

Conclusion:

Nowadays, e-contracts are the most widely used technology. The use of credit or debit cards and internet banking, all of which are on the rise, as well as educated and proficient computer users, will assist this increase even more. The contract must address every aspect, starting with payment and ending with delivery. Such law will aid in expanding the e-contract's reach and limiting websites that grow for a short while before going dark owing to a lack of adequate funding. The key fundamental connection is the trust of the customers, which should be engaged at all costs, and legislation in this field will discover the crooks who have utilized the internet without any direct physical intervention.[17]

In addition, based on the current situation, it is absolutely correct to say that the proliferation of COVID-19 is having a substantial impact on everyone of

[14] Available at https://www.indialawoffices.com/legal-articles/e-contracts-and-validity-india

[15] Available at https://www.nolo.com/legal-encyclopedia/electronic-signatures-online-contracts-29495.html

[16] Available at https://www.mondaq.com/india/contracts-and-commercial-law/908604/e-contract-in-times-of-covid-19

[17] Available at https://www.fieldfisher.com/en/insights/a-guide-to-electronic-signatures-during-coronavirus

our daily lives. A number of challenges arise when contracts, deeds, and other important papers are signed and executed because of the continued social isolation and self-isolation caused by the fact that a sizable portion of the workforce now works from home. People might not be able to physically sign documents and may not have access to printers, scanners, or both. It may also be very difficult to witness signatures. Considering that it is doubtful that travel limitations will be relaxed very soon, businesses and individuals will need to adjust and find substitute ways to sign documents.[18]

[18] Available at https://shodhganga.inflibnet.ac.in/bitstream/10603/107814/12/12_chapter%20v.pdf

ARTIFICIAL INTELLIGENCE (AI) AND JUSTICE SYSTEM.

Suhas Narhari Toradmal[1] And Dr.Nayana Nitin Mahajan[2]

Abstract

The integration of Artificial Intelligence (AI) into the criminal justice system in India represents a significant paradigm shift with far-reaching implications. This article explores the multifaceted intersection of AI and the criminal justice system in the Indian context, investigating the potential benefits and challenges associated with the deployment of AI technologies. The article begins by providing an overview of the key areas where AI applications can enhance efficiency, accuracy, and fairness and identifies current state of the criminal justice system in India. It delves into the utilization of AI in crime prevention, investigation, and adjudication processes, highlighting the promising outcomes and improvements witnessed in these domains. With every progressing day Artificial Intelligence (AI) is getting a grip in every sphere of our daily life. In such an environment, Criminal Justice System, which is an ever-expanding domain trying to cater to the contemporary need of the society to make it a safer place to live, it tries to adopt every possible method and technique to accomplish this objective. AI too has been adopted in the functioning of the criminal justice system to adopt more scientific and sophisticated approach to crime prevention and crime detection. Artificial intelligence (AI) has been making waves across various industries worldwide, and the criminal justice system is no exception.

Keyword : Public Prosecutor, Artificial intelligence, Criminal justice system

[1] Research Scholar, KBC, NMU Jalgaon.

[2] Principal, Dr.Ulhas Patil Law College, Jalgaon.

Introduction:

The integration of Artificial Intelligence into the role of Public Prosecutors has the potential to revolutionize the justice system. While AI is not a replacement for human judgment, It can be act as powerful tool to support public prosecutors in performing their duties more efficiently and effectively.AI enhances efficiency, accuracy and fairness in legal proceedings by automating routine tasks assisting legal research and predicting case outcomes[3].

Evidence Processing and Investigation support:
AI can process vast amounts of digital evidence, including CCTV footage, phone records, emails and forensic reports.

Machine learning algorithms: It help identify patterns in evidence, linking suspects to crimes more efficiently.AI powered tools like facial recognition and voice analysis assist in verifying identities and detecting fraudulent statements.

Framing of Charges and Legal Issues: AI helping drafting charge sheets by identifying the most relevant legal provisions.AI-assisted tools can ensure consistent application of legal standards, reducing errors in framing charges[4]. The Public Prosecutor plays a pivotal role in the Indian criminal justice system. As representatives of the state, they are responsible for prosecuting criminal cases on behalf of the government. Their primary duty is to ensure that justice is served while maintaining the rights of the accused. The integrity of a person chosen to be in charge of a prosecution does not need to be emphasised. The purpose of a criminal trial being to determine the guilt or innocence of the accused person, the duty of a Public Prosecutor is not to

[3] www.indiankanoon.com

[4] Artificial Intelligence Law. Swan, Edward J.

represent any particular party, but the State[5]. The prosecution of accused persons has to be conducted with the utmost fairness. In undertaking the prosecution, the State is not actuated by any motives of revenge but seeks only to protect the community.

There should not therefore be "an unseemly eagerness for, or grasping at a conviction". A Public Prosecutor should be personally indifferent to the result of the case. His duty should consist only in placing all the available evidence irrespective of the fact whether it goes against the accused or helps him, before the court, in order to aid the court in discovering the truth. It would thus be seen, that in the machinery of justice, a public prosecutor has to play a very responsible role: the impartiality of his conduct is as vital as the impartiality of the court itself.

He is the holder of public office and has to discharge his function impartially and independently. They are not to suppress material facts, evidences or witnesses even if it is not in favour of his case. In other words, he is not to secure a conviction by any means. His duty is to aid the court in finding the truth.

Legal Provisions under BNSS Regarding public Prosecutor

Section 18: Public Prosecutors

The Bharatiya Nagarik Suraksha Sanhita (BNSS) outlines the provisions for Public Prosecutors, including their appointment, eligibility, and roles. Section 18 of the BNSS focuses on the appointment of Public Prosecutors and Additional Public Prosecutors for High Courts and districts, requiring consultation with the relevant High Court and District Magistrate/Sessions

[5] Bhartiya Nagari Surkasha Sanhita-2023

Judge for appointments. Section 19 details the appointment and functions of Assistant Public Prosecutors.

Public Prosecutor plays a vital role in the administration of the Criminal Justice System in India. The Police, the Prosecution, the Judiciary, and the Prison and Correctional Services constitute the components of the Criminal Justice System.

A Public Prosecutor represents the State in a criminal case as a crime is considered to be an offense against the State. His role came into the picture after charge sheet is submitted by the police. He presents the multiple aspects of the case during the trial and assists the Court in the delivery of Justice. Accordingly guidelines on the Role of the Prosecutors requires Prosecutors to perform their duties fairly, impartially, and consistently, protecting human dignity, upholding human rights and avoiding all political, social, religious, cultural, sexual or any other kind of discrimination[6].

In order to ensure the fairness and effectiveness of prosecution, prosecutors must strive to cooperate with the police, the courts, the legal profession, public defenders and other government agencies or institutions. Complying with these provisions, the Code of Criminal Procedure, 1973 mandates the appointment of Public Prosecutors in the High Courts and District Courts.

Public Prosecutor and His Appointment

A Public Prosecutor is a lawyer who acts for the government against someone accused of a crime and prosecutes them in the Court.

[6] Bhartiya Sakshya Adhiniyam-Dr.Ashok Kumar Jain, Acent Publication

According to the Code of Criminal Procedure, 1973 (Cr.PC, 1973) "Public Prosecutor" means any person appointed under Section 24, and includes any person acting under the directions of a Public Prosecutor. While, under Bhartiya Nagarik Suraksha Sanhita, 2023 (BNSS), he is a person appointed under Section 18 and includes any person acting under his directions[7].

Categories of Public Prosecutor

Section 24 of Cr.PC (now, Section 18 of BNSS, 2023) provides three main categories of Public Prosecutors:

A. Category I: Those who are attached to a particular High Court or district;

B. Category II: Assistant Public Prosecutors who are linked to a particular case or a class of cases but in a specified jurisdiction;

C. Category III: Special Public Prosecutors appointed under Section 24(8) of the Code [or Section 18(8), BNSS].

A. Appointment of Public Prosecutor under Cr.PC, 1973 (Category I)

Section 24 of the CrPC, 1973 (Section 18 of BNSS, 2023) mandates the appointment of a Public Prosecutor or Additional Public Prosecutor for every High Court and District Court. The provisions of this section can be explained in the following steps:

1. Who Appoints

Section 24(1) of Cr.PC provides that the Central Government and the State Government appoint a Public Prosecutor and may appoint one or more Additional Public Prosecutors for every High Court and they have to conduct

[7] www.indiankanoon.com

prosecution, appeal or other proceeding in such Court on behalf of the respective Government.

This provision is similar to Section 18(1) of BNSS. However, the BNSS includes a proviso clause for NCT Delhi, according to which only Central Government shall appoint the Public Prosecutor or Additional Public Prosecutors, after consultation with the High Court of Delhi.

The point to be noted here is that the power of the State Government to appoint a Public Prosecutor or Additional Public Prosecutor would extend only for conducting prosecution, appeal or other proceedings in the courts within that State *(Jayendra Saraswathi Swamigal v. State of Tamil Nadu*[8]*).*

2. Panel of Names

To appoint a Public Prosecutor or Additional Public Prosecutors for every district, the District Magistrate u/s 24(4) of Cr.PC (now, Section 18(4) of BNSS) shall prepare a panel of names in consultation with the session judge. The panel of names includes the name of those who are in their opinion fit to be appointed as Public Prosecutor or Additional Public Prosecutor. The State government shall appoint them from these panel of names u/s 24(3) Cr.PC [similar to Section 18(3) BNSS].

Further, Section 24(5) Cr.PC [similar to Section 18(5), BNSS] puts an obligation on the State that it cannot appoint any person as the Public Prosecutor or Additional Public Prosecutor unless his name is not in the list of *panel of names* prepared by the District Magistrate.

[8] AIR 2005 SC716

3. Regular Cadre of Prosecuting Officers

Section 24(6) of the Cr.PC mandates that, if there exists a regular Cadre of Prosecuting Officers in a State, the State Governments are obliged to appoint Public Prosecutors or Additional Public Prosecutors only from amongst the persons constituting such cadres. The same provision is also laid down by BNSS u/s 18(6).

If in the opinion of the State Government, no suitable person is available in such Cadre for such appointment that Government may appoint a person as Public Prosecutor or Additional Public Prosecutor, as the case may be, from the panel of names prepared by the District Magistrate under sub-section.

The expression "regular cadre of Prosecuting Officers[9] "as held by Supreme Court in the case of *A.K. Ahlawat v. State of Haryana, (2010)*[10] comprised a service with the Assistant Public Prosecutor at the lowest level and Public Prosecutors at the top. Further the court in the case of *K.J. John, Asst. Public Prosecutor v. State of Kerala* held that in case a regular cadre of Prosecuting Officers did not go up to the Public Prosecutor at the top, the State Government cannot be considered as bound to appoint Public Prosecutor or Additional Public Prosecutor only from among the persons constituting such cadre under the Code of Criminal Procedure for conducting cases in the sessions court.

4. Experience

Section 24(7) of Cr.PC [similar to Section 18(7) of BNSS] laid down a mandatory provision that the person has been in practice as an advocate for not less than seven years to be appointed as an Additional Public Prosecutor or an Additional Public Prosecutor under sub-section (1) or sub-section (2) or sub-section (3) or sub-section (6) of this Code.

[9] Code of Criminal Procedure1973

[10] (2010) SCC On Line P&H 5139

In short, Public Prosecutor and Assistant Public Prosecutor is appointed by Centre and State Government in High Court and in a district the Public Prosecutor or assistant Public Prosecutor is appointed from the panel of names prepared by District Magistrate in general.

B. Assistant Public Prosecutor (Category II)

Section 25 of Cr.PC [S.19 of BNSS] deals with the appointment of Assistant Public Prosecutors in the district for prosecution in the courts of Magistrate. Section 25(1) of Cr.PC states that the State Government shall appoint in every district one or more Assistant Public Prosecutors for conducting prosecutions in the Courts of Magistrates.

Further, u/s 25(2) Cr.PC, no police officer shall be eligible to be appointed as an Additional Public Prosecutor subject to the provision under sub-section (3). According to Section 25(3), where no Assistant {Public Prosecutor is available for the purpose of any particular case, the District Magistrate may appoint any other person to be Assistant Public Prosecutor in charge of that case:

Provided that a police officer shall not be so appointed –

(a) If he has taken any part in the investigation into the offence with respect to which the accused is being prosecuted; or

(b) If he is below the rank of inspector.

C. Special Public Prosecutor

The Central Government or the State Government may appoint, for the purposes of any case or class of cases, a person who has been in practice as an advocate for not less than ten years as a Special Public Prosecutor u/s 24(8) of Cr.PC[s.18(8) of BNSS].

It is to be noted that Supreme Court in the case of *Poonamchand Jain v. State of M.P*[11]. held that the appointment of Special Public Prosecutor is not with reference to the High Court or district, but is an appointment for a case or class of cases in any court. Without disclosing a special reason, order appointing a Special Public Prosecutor is illegal.

Key Roles of a Public Prosecutor

A public prosecutor is responsible for

1. Representing the State in criminal cases
2. Ensuring Justice: Their role is not just to secure convictions but to ensure fair and impartial trials
3. Analyzing Evidence: They must examine evidence provided by law enforcement and decide whether to proceed with charges
4. Presenting Arguments: Prosecutors must take strong evidence based argument in court.
5. Guiding Investigation: Collaborating with law enforcement agencies to strengthen cases.
6.

Role of AI in supporting Public Prosecutors:

1. Enhancing Evidence Analysis: AI can analyze large volume of evidence, such as videos, text and digital data, identifying relevant patterns or inconsistencies.
2. Cross-referencing: AI tools can cross check evidence with prior cases, highlighting similarities or contradictions.
3. Legal Research and Documentation: Quick Access to Legal Precedents-AI system like legal research tools can identify relevant case law and precedents in seconds.

[11] (2001) SCC OnLine MP27

4. Drafting Documents: AI can assist in preparing charges, legal notices and reports, reducing the workload on prosecutors.
5. Case Prioritization: Resource management: AI can help prioritize cases based on their complexity societal impact or likelihood of conviction.
6. Predictive Analytics: Using historical data,AI can predict case outcomes, helping prosecutors decide whether to proceed or seek plea agreement.
7. Fairness and Bias Detection: Identifying Disparities: AI can detect patterns of bias in case handling, such as disparities in charging decisions or sentencing recommendations, ensuring fairness.
8. Reducing Human Error: AI helps minimize subjective errors in evidence analysis or legal reasoning.
9. Collaboration with Law Enforcement: AI-powered tools can streamline the communication between law enforcement and public prosecutors ensuring better case preparation and stronger investigation.
10. Sentencing and Plea Bargain Recommendations: AI can provide data-driven sentencing guidelines or plea bargain suggestions based on historical data, ensuring consistency.
11.

Challenges in using AI for Public Prosecutors:

1. Ethical Concerns: AI decisions must not replace human judgment, prosecutors must refrain ultimate authority.
2. Bias in AI Algorithms: AI systems must be transparent and free from inherent biases to ensure justice.
3. Data Privacy: Handling sensitive information requires robust safeguards
4. Accountability: Prosecutors must ensure AI tools are used responsibly and explain how AI influences their decisions.

In Prabhu Dayal Gupta V. State the High Court Delhi[12] Viewed that a Public Prosecutor must be free & flaxen in any presentation of the criminal case. He has to present a whole movie of the prosecution case. This is a duty of the public prosecutor to present an absolute picture and not the one-sided movie. He must be impartial and not be biased to the state (hearing) or bias to the accused side. A public prosecutor has a duty to be fair to both sides of the criminal case[13].

The job and role of a Public Prosecutor in between the pre-trial stage of a criminal case are deciding. The role of a public prosecutor is crucial in this stage. It is important in following ways; to deciding in the degree of discretion in the choices of "not to charge" and "to charge", in the choices of control and the degree of coordination between the prosecutor and police, the space of the prosecutor to facilitate the investigators or the investigation, to obtain a feasible prosecution file, on the need to extend the work of the Public Prosecutor by attributing to him one or more powers which are presently practiced by the Magistrate to awaken the strict dismissals.

All these above-said functions are to be performed in this stage of the prosecution. The cornerstone of the play includes standards: prosecutions will be initiated just when there is sufficient proof and relevant evidence not otherwise; the choice to induct charge or not to introduce charge is a choice of prosecution and should not be the jurisdictional capacity of the court. These powers can be transferred to the public prosecutor for justice and to strengthen the criminal administration of justice.

Conclusion:

[12] Bhartiya Nyaya Suraksha Sanhita,2023

[13] Radha Krishna krupa sagar,The role of public prosecutor in criminal justice system 174-175(2013)

AI can serve as a valuable assistant to public prosecutors, enhancing their ability to process information, reduce backlogs and promote fairness in the justice system. However, its use must be balanced with ethical considerations and human oversight to ensure justice is upheld in its true sense.

ARTIFICIAL INTELLIGENCE IN INVESTIGATION & JUDICIARY: A GLOBAL PERSPECTIVE

Muskan Rani[1]
Dr. Jayashree Khandare[2]

Abstract

The application of Artificial Intelligence is on the surge and has both direct and indirect implications on the future of justice across the globe. This dissertation research has sought to consider the current application and traditional use of AI developing science and technology in decision-making, law drafting, crime scene formulation investigations, and management of the evidential material and research across law. It seeks to consider the consequential benefits against the ethical underpinnings of algorithmic bias and data privacy and their respective possible risks. It is valuable because the distinct probability view of human oversight will change with possible risks. This research has given the future of artificial justice a deserving plausible future. Through cross-bordering comparative analyses on multiple published discussions on the various judiciary systems across the globe, this research aims to help countries develop a system framework for integrated responsible futurist, and the role of AI will play across the International justice systems.

Introduction

Artificial intelligence is reshaping various sectors globally, and criminal justice and investigative industries are not exempt from its impact. AI has the

[1] Author, Jindal Global Law School, OP Jindal Global University.

[2] Correspondence Author & Assistant Professor, Bharti Vidyapeeth Deemed to be University, New Law College, Pune

power to revolutionize the legal business if it is allowed to help in making decisions with greater accuracy, enhancing the process, and increasing exclusion-based rulings. Specifically, AI can perforce process, raise evidence-based decision-making, and enable scrutiny of responsibility and openness in the court and investigation settings. The judiciary and investigators grapple with the customary challenges of high filing, insufficient budgets, and the need to address the changing forms of criminality. These factors have combined to create other daunting aspects such as high levels of impunity, lack of victim orientation, pre-trial detention, non-use of custodial sentences, and backlogged justice.[3]

The primary research question of this paper is: What benefits and drawbacks does AI bring to investigating and judicial processes worldwide? This study aims to investigate the impact of AI on legal processes, possible applications of AI to traditional issues, and possible moral and legal consequences of its use. Search results have made clear the possible benefits of AI in the legal sector. Among these benefits include document review, legal research, and prompt, direct responses to legal queries. They also raise concerns with privacy, security, and the potential for biases and errors in citations and AI research.[4]

Moreover, when AI has been used in the judicial system, problems with fairness, openness, and bias have emerged. Furthermore, required are well defined legal frameworks and international cooperation. The legal and ethical frameworks controlling the application of AI in the legal system, the historical evolution and current state of AI applications in the investigative

[3] United Nations Office on Drugs and Crime (UNODC). Integrated approaches to challenges facing the criminal justice system.<https://www.unodc.org/> accessed 28 March 2024

[4] Esfandi G, “The Potential and Drawbacks Of Using Artificial Intelligence In Legal Field Plaintiff Magazine <https://plaintiffmagazine.com/recent-issues/item/the-potential-and-drawbacks-of-using-artificial-intelligence-in-the-legal-field> accessed 28 March” 2024

and judicial systems, and the opinions of different nations on the incorporation of AI in the legal sector will all be examined in this study.[5]

Chapter 2: Applications of AI in Investigations

AI plays a crucial role in various aspects of crime scene analysis, evidence processing, predictive policing, and proactive crime prevention. Researchers and scholars have highlighted the significant impact of AI in enhancing forensic science and criminal investigations. AI technologies are aiding forensic experts and investigators in formulating logical evidence, reconstructing crime scenes in 3D, and handling evidence efficiently.[6]

Moreover, AI is being utilized in public safety video and image analysis, DNA analysis, gunshot detection, and crime detection to address criminal justice needs effectively[7]. AI algorithms have the potential to detect crimes that might otherwise go undetected, ensuring greater public safety by investigating potential criminal activity and increasing community confidence in law enforcement along criminal justice system.

Furthermore, AI technology is revolutionizing crime scene investigations by providing precise and objective analysis of risks posed by sentenced individuals and enhancing bloodstain pattern analysis through advanced segmentation techniques[8]. AI algorithms are capable of predicting criminal behavior, identifying high-risk individuals, disrupting criminal enterprises, and assisting law enforcement professionals in safeguarding the public in innovative ways.

[5] Reiling, A. D. (Dory). "Courts and Artificial Intelligence" (2020) 11(2) “International Journal for Court Administration accessed 28 March” 2024

[6] Ekta “B Jadhav, Mahipal Singh Sankhla and Rajeev Kumar, 'Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation' (2020) 15(8) Seybold Report” 2064

[7] Christopher “Rigano, Using Artificial Intelligence to Address Criminal Justice Needs (2021) 22 J Crim” L 123

[8] Pratima Gund, “Investigating Crime: A Role of Artificial Intelligence in Criminal Justice” (June 2023) https://www.researchgate.net/publication/371415843_INVESTIGATING_CRIME_A_ROLE_OF_ARTIFICIAL_INTELLIGENCE_IN_CRIMINAL_JUSTICE accessed 12 May 2024

In essence, the scholarly view on AI in crime scene analysis and criminal justice emphasizes its transformative potential in improving evidence processing, crime detection, and proactive crime prevention. By leveraging AI technologies, law enforcement agencies can enhance their investigative capabilities, ensure public safety, and contribute to more effective and efficient crime prevention strategies.

Scholar View on Applications of AI in Investigations

Ryan Calo, a professor at the University of Washington School of Law, is one of the contributors to Hannity. Calo is credited for conducting research on the potential of AI solutions within the sphere of AI and, in particular, law enforcement. His studies address how AI can become instrumental in multiple functions of the investigation, such as processing extensive data and detecting commonalities or suggesting more probable leads to law enforcement. [9]

Calo draws on the above to write that AI has an outstanding capacity to modernize analysis at crime scenes, data processing, and criminal prevention actions. However, he asserts that AI investigations should comply with the utmost legal and ethical standards. In particular, Calo discusses the importance of protection of individual rights and privacy in case AI systems are not regulated closely enough. [10]

Calo further points out that the legality of use of AI in criminal justice will depend on level of transparency and responsibility. Apart from the legal implications in use of AI in a trial, the legal system might need to tackle problem of double duty of an instrument and a person’s identity. Namely, AI may “become” a person. As a result, understanding of being used in “false

[9] Ryan Calo, “Robotics and Lessons of Cyberlaw (2015) University of Washington School of Law Digital Commons
https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=1022&context=faculty-articles accessed 12 May 2024”

[10] María P Angel and Ryan Calo, 'Distinguishing Privacy Law: A Critique of Privacy as Social Taxonomy

belief" in robots would lead to irregular or uncertain judgements for uncertain or unreliable results.[11]
Another example of Calo's investigation into the societal consequences of AI within the criminal justice field concerns the effect on official's decision making in policing. The author establishes that AI systems, and specifically humanoid robots, elicit different reactions and decision processes in comparison with traditional law enforcement systems. Summarily, Ryan Calo's work highlights the enormous possibility of effective AI enforcement techniques and the necessity of rigorous ethical and legal measures to regulate this potential. For his part, arguing for a middle ground between utilizing AI's full potential while simultaneously guarding against its drawbacks is Calo adding to the ongoing debate on AI's responsible development and use in criminal investigations.

He is one of the advocates of the view considered above. When they speak of this position, judgements of legal scholar and ethicist Frank Pasquale are of particular importance. His research was crucial for starting the conversation on AI's ethical issues and its utilization as an investigatory instrument in criminal cases. In particular, Pasquale raises the issues of the question of transparency, responsible behavior, and fairness when using it. He explicitly formulates his attitude towards the question: Algorithms of repression must be subject to public scrutiny and public veto. In other words, the author advocates for significant internal control over the application of AI in the police in order to avoid abuses of authority and discrimination against the public. Moreover, Pasquale points out the necessity of what he calls

[11] Ryan Calo, 'Robots in American Law' <http://euro.ecom.cmu.edu/program/law/08-732/AI/Calo.pdf> accessed 12 May 2024

algorithmic transparency in other words, the public should know how the system works and what exactly it does.[12]

Pasquale similarly calls for safeguarding fundamental human rights and values. It is in advocating for a human-based approach to AI deployment in investigations. Specifically, he proposes that the fundamental notions of justice, equality, intra-human respect, faithfulness to the innocent, and freedom should prevail in terms of AI use. In addition to that, Pasquale also underscores the challenge of ensuring accountability for the use of AI. He recommends creating clear and diverse tracks of accountability that would enable to people and organizations to take the blame for their AI systems dealers.

Apart from his research work, Pasquale manifested vibrant involvement in the policy-making discussion around the necessity of introducing the AI regulation in the field of the criminal justice system. He articulates the idea of the formation of the moral principles and rules that would set the boundaries of AI application in the arena of maintenance of law, as well as the necessity of a perspective aimed to make it viable through the cooperation of the scientists, legislators, and civil society. Hence, the conversation addressing the moral aspect of the application of artificial intelligence technologies in criminal investigations was highly influenced by Frank Pasquale. Plus, this research has promoted the establishment of the integrations of the AI into the practice of the law enforcement agencies that is more responsible and ethical due to the open, fair, and responsible perspective.

Crime Scene Analysis

[12] European Parliament Research Service, AI: From Ethics to Policy (2020) <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU%282020%29641507_EN.pdf> accessed 12 May 2024

Facial Recognition: The area of my political research would be transformed by AI-powered face recognition. In pictures and videos, AI can scan people's faces, which it could then speedily match against databases of criminals or missing individuals. Not only would this aspect assist identify people at risk, but it would also help capture miscreants. Nonetheless, there are difficulties with accuracy, particularly when the angle or height of the face is varied and lighting is unclear. One possible negative effect might be biased or unjust facial recognition as a result of biased training results. However, privacy concerns may arise as the employment of face recognition technologies expands. Regardless of the issues, facial recognition AI technology is a beneficial general-purpose tool for scientists; however, the research and expansion of this technology must be approached from an ethical and legal standpoint.[13] AI-powered face recognition is inspiring a new area of analysis. With the help of pictures and videos, AI may be used to disappear suppresses or missing people by comparing faces against its enormous databases of surefire suspects. Perhaps it may not be enough to stop outlaws, but it is more than sufficient to find a missing child or elder. Accuracy is an issue, particularly when appearances or lighting differ. Additionally, there is a risk of injustice due to the incomplete identification of training data. Furthermore, the increasing use of face recognition technology raises privacy concerns.

Despite those challenges, AI-powered facial recognition remains a valuable element of efficient investigation capabilities. Nevertheless, ethical and legal aspects of the widespread implementation and further development of the technology should be taken into account to a greater extent. The more AI-powered facial recognition progresses, the more regulations and definitions need to be applied to control its implementation. Policymakers, law enforcement, and tech companies must cooperate to ensure that the privacy,

[13] AI in Criminal Justice (Master's in AI) <https://www.mastersinai.org/industries/criminal-justice/> accessed 28 March 2024

bias, and accuracy issues are addressed while enabling the full utilization of the potential of AI-powered facial recognition in assisting police rescue crime victims.

As much as AI-powered facial recognition is being developed, there is also the need for continuous research and testing to enhance its accuracy and reduce potential bias. For this use of virtual reality, it is essential for the group consisting of researchers, developers and end-users to check detect and repair any issues likely to happen before they affect anyone virtually. Finally, the maximum benefits of AI-powered facial recognition and the limited potential harm that can happen to individuals using it can be ensured when transparency and accountability and ethical values are considered.

Fingerprint Analysis: Fingerprinting is one of the oldest forms of churning out criminal records, and thanks to AI, this form of forensic science has been revolutionized. Fingerprinting has also gone the way of automation, and the software we now use can compare and print vast databases in a matter of seconds. This has left specialists with more time to pursue complex cases and has reduced the risk of crimes going unsolved for more than a few hours or less.

Fingerprint analysis, a cornerstone of forensic investigations, is undergoing a revolution fueled by AI. Traditionally, a labor-intensive process involving manual comparisons with databases, fingerprint analysis is being transformed by AI's ability to:

- Expedite Matching: AI algorithms can analyze massive fingerprint databases in seconds, freeing up examiner time for complex cases and accelerating suspect identification. This aligns with the efficiency gains highlighted in the travel and tourism industry regarding AI-driven facial recognition.

- Enhance Accuracy: AI can assist in analyzing faint or partial prints that might be missed by humans, similar to how some facial recognition systems are being explored for CCTV footage analysis. Additionally, AI can identify subtle fingerprint variations, potentially holding valuable clues, as highlighted in the scholar work you provided.
- Uncover New Evidence: AI's ability to analyze vast datasets and detect intricate details surpasses human capabilities. This opens doors for uncovering crucial evidence in complex investigations.

However, challenges remain:

- Bias in AI Systems: As with facial recognition, AI systems in fingerprint analysis can inherit biases from training data, leading to misidentifications. Rigorous testing and diverse datasets are crucial to mitigate this, as emphasized in the scholar work.

The Future of Fingerprint Analysis with AI

Despite these hurdles, the potential of AI in fingerprint analysis is immense. By harnessing its capabilities, forensic investigations can benefit from:

- Faster Processing: AI can expedite fingerprint analysis, freeing up human expertise for crucial tasks.
- Improved Accuracy: AI can assist in identifying faint prints and subtle variations, potentially leading to more accurate results.
- New Investigative Avenues: AI's ability to analyze vast datasets may uncover hidden connections and evidence, aiding in solving complex cases.

Evidence Processing

Analyzing Large Datasets: The ocean of data created today, ranging from phone records to social media posts, is simply too much for traditional investigation techniques to process. AI comes into play here. Because it "interrogates" these huge data sets, AI is a brilliant detective's assistant. It is capable of discovering subtle patterns and relationships that the human eye might overlook. Sure, maybe the suspect's phone ping lives a few blocks from the criminal scene to the victim's social media post. Or maybe the criminal ring is linked by hundreds of thousands of minor financial transactions. Such capabilities to uncover hiding connections streamline cases and speeds up bringing criminals to court. Allowing AI to investigate, however, is determined by the likelihood and level of the data analyzed orchestration.[14]

Benefits of AI in Analyzing Large Datasets:

- Faster Processing: AI can analyze massive amounts of data (phone records, social media) much quicker than humans, identifying patterns and connections that might be missed.
- 24/7 Work: Unlike humans, AI can work tirelessly, ensuring a thorough investigation.
- Reduced Bias: AI algorithms can be designed to be objective, potentially reducing unconscious bias that can affect human investigators.

Challenges and Considerations:

- Data Quality: The accuracy of AI insights depends heavily on the quality and diversity of the data used to train it. Biased or incomplete data can lead to misleading results.

[14] Lunter J, "Can criminal investigations rely on AI? (24 November 2023) Biometric Update <https://www.biometricupdate.com/202311/can-criminal-investigations-rely-on-ai> accessed 28 March" 2024.

- Ethical Use: It's crucial to ensure AI is used responsibly and ethically. Law enforcement needs to carefully manage data used to train AI systems to avoid bias and ensure fairness.
- Overall, AI has the potential to be a powerful tool for criminal investigations, but it's important to be aware of the limitations and use it responsibly.

.

Gunshot Recognition: Gunshot Recognition: AI algorithms play a crucial role in differentiating gunshot patterns, determining timings, and estimating identifying probabilities, thereby enhancing evidence analysis.[15] As evidenced in different studies, these algorithms can achieve high accuracy in identifying gunshot sounds, down to the gun's location and type in actual-time situations. Using Deep Learning techniques and complex methods such as audio to spectrogram image conversion and utilizing CNN architectures, AI systems can accurately process gunshot sound recordings by distinguishing between various gun models at incredibly high accuracy levels. The additional application of pre-trained models and transfer learning also significantly improves AI system performance such that it is capable of processing large bulks of data, thus enhancing recognition levels.

In summary, the gunshot recognition technology powered by AI is not only instrumental in generating data on gunshot sounds but is also pivotal in the creation of integral forensic findings surrounding the event, including the nature of the firearm, the location of fire, and the timing of the incident. The results derived from the utilization of these algorithms are game changers in the analysis of evidence, providing invaluable assistance to law enforcement in situational knowledge, better response time, and fostering a community that is free from any risks.

[15] Purdue University Northwest, "PNW Computer Science research team tests AI-powered gunshot detection technology (19 January 2024) <https://www.pnw.edu/pnw-computer-science-research-team-tests-ai-powered-gunshot-detection-technology> accessed" 15 May 2024

Predictive Policing and Proactive Crime Prevention

AI and machine learning techniques are being increasingly used by law enforcement and financial institutions to predict and detect criminal activities like fraud. Here are some key ways these technologies are being applied:

> Crime forecasting: AI algorithms can analyze large volumes of data on crime, social factors, and legal precedents to predict potential criminal activities and identify people at risk of being involved in crime. Machine learning models can detect anomalous patterns that may indicate criminal enterprises.
>
> Fraud detection: AI and ML are widely used in payment fraud detection. Techniques like random forest, decision trees, neural networks and deep learning are used to learn from transaction data and spot fraudulent activity in real-time[16].AI-based fraud detection is significantly faster and more accurate than traditional rules-based systems.
>
> Predictive policing: AI can analyze data from various sources like crime reports, social media, and surveillance cameras to predict where and when crimes are likely to occur. This allows law enforcement to proactively deploy resources to prevent crimes.
>
> DNA analysis: AI is being used in forensic DNA analysis to identify suspects and establish cause of death. Machine learning algorithms can analyze complex DNA mixtures much faster than humans.

[16] Stijn Boon, "How AI and ML are used in payment fraud detection (16 use cases)" (Nomentia, 25 March 2022) <https://www.nomentia.com/blog/ai-machine-learning-in-fraud-detection accessed> 15 May 2024.

Despite the many potential benefits of AI in the fight against crime, the technologies raise concerns of security and privacy, bias, and lack of transparency as AI systems are getting more and more sophisticated. Nevertheless, AI is a potent instrument that can greatly improve the work of both law enforcement agencies and financial corporations in the fields of criminal detection and prevention.

Proactive Measures: AI technologies are indeed transforming law enforcement by enabling proactive measures that shift focus from reactive to proactive strategies. By analyzing crime data, identifying patterns, and optimizing resource allocation, AI can significantly reduce criminal activities. Here's how AI facilitates this shift:

- Analyzing Crime Data: AI algorithms can process vast amounts of historical crime data to identify patterns and trends that may indicate potential criminal activities[17]. By analyzing this data, law enforcement agencies can gain insights into where and when crimes are likely to occur.
- Identifying Patterns: AI-driven predictive policing can identify patterns in crime data that may not be immediately apparent to human analysts. This allows law enforcement to anticipate criminal activities and take preventive action[18].
- Optimizing Resource Allocation: AI helps police departments allocate their resources efficiently by guiding them on where to deploy officers based on predictive insights. This proactive approach ensures that law

[17] Ripla A, 'Predictive Policing and Crime Prevention' (LinkedIn, 19 May 2024) <https://www.linkedin.com/pulse/predictive-policing-crime-prevention-andre-ripla-pgcert-91xbe/> accessed 15 May 2024.

[18] "Predictive Policing using Machine Learning (With Examples) <https://www.cogentinfo.com/resources/predictive-policing-using-machine-learning-with-examples">

enforcement is present in high-risk areas, deterring criminal activities before they happen.

By leveraging AI for proactive measures, law enforcement agencies can enhance public safety, prevent crimes before they occur, and optimize their operations to effectively combat criminal activities.

Chapter 3: Applications of AI in the Judiciary

AI is revolutionizing the judicial system, offering innovative tools and capabilities that enhance legal research, decision-making, and court processes.

AI-Powered Legal Research Tools

AI-driven legal research tools have significantly impacted case preparation by streamlining the process of analyzing vast legal data. These tools enable lawyers to identify precedents, streamline clerical tasks, and support judges in predicting outcomes like criminal sentence duration and risk assessment recidivism scores.[19] Despite the potential of AI in befitting legal research, the technology brings up a number of demographic concerns regarding ethical inaccuracies issues affecting the content. Given these realities, the safeguards designed to prevent leaks of voter information can actually heighten the risks posed by AI's methodological biases. Hence, it is the ethical responsibility of the lawyers to understand how AI is incongruous.[20] Relying on AI-generated work products does not ensure the ethical duty of professional competence. This is because a lawyer is not deemed to understand the AI system if he or

[19] JTC Resource Bulletin, "Introduction to AI for Courts <https://www.ncsc.org/__data/assets/pdf_file/0027/98910/JTC-AI-paper-update-3.5.24.pdf> accessed 28 March" 2024.

[20] Cerny, J., "Delchin, S., & Nguyen, H. Legal Ethics in the Use of Artificial Intelligence. (2019) Available at: <https://www.squirepattonboggs.com/-/media/files/insights/publications/2019/02/legal-ethics-in-the-use-of-artificial-intelligence/legalethics_feb2019.pdf > Accessed" 19 May 2024

she does not understand how the AI operates and does not actively oversee any work product an AI system may generate. Moreover, work product and conclusions reached by AI are not substitutes for human judgment and must be reviewed for accuracy, regulatory compliance, and completeness.[21]

In AI, bias is an issue resulting from the fact that AI tools are trained to use information rather than how they are programmed. Therefore, if the information on which the AI is trained is biased, then biased outcomes are expected. It is critical for lawyers involved in the use of AI to understand how bias can affect AI results. The bias in the use of AI can create vulnerabilities for the clients represented by the lawyers. It is important to critically examine the AI work products and the kinds of biases that may be present. This should be done before submitting the work products to inform the cases. Lawyers have an ethical obligation of being competent professionals in using technology in their practice. It is important to understand that using AI as a tool in making conclusions lacks due diligence since not much is known about the capacity and application of that technology in the legal profession and society in general. Lawyers will be responsible for how they use the knowledge from AI to make informed decisions on legal matters. Therefore, it is important to assess the possible bias in the algorithm used in drive the decision-making tools.

The use of AI in legal research poses an ethical dilemma regarding the use of AI in the legal system. Although AI tools can aid legal research and decision-making, lawyers are ethically obligated to provide competent legal representation and cannot abdicate this obligation to AI. Lawyers must

[21] "Duplantier, T., 'AI and Ethical Concerns for Legal Practitioners' (January 08, 2024) <https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/ai-and-ethical-concerns-for-legal-practitioners">

carefully check and evaluate the accuracy and potential biases in AI-generated work product before submitting it in legal settings.

AI's Role in Decision-Making

AI has huge potential in assisting the judges in decision-making processes. Access to AI predictive analytics will enable the judge to make informed decisions, assess risks, enhance the efficiency of the judicial system and improvement of sentence guidelines. As a self-learning tool, the AI should function under the transparency, accountability, and human oversight to avoid unintended consequences. One of the ways in which AI could provide on-judged assistance is by analyzing the enormous volume of case laws and precedence. AI will scan and identification of count cases, the relevant legal principle and history of sentencing, will enable the judges to be consistent and well-informed in their decisions and complying with the law, thus increasing fairness and equity.[22]

AI-powered risk assessment tools can support judges in forecasting the chances of a defendant's recidivism or failure to attend trial. Considering the criminal history, social determinants, and behavior patterns, such tools can be helpful for judges in making the most appropriate decisions on granting bail, imposing sentences, or prescribing reintegration programs. However, it is crucial to audit these tools regularly to detect and eliminate any bias and to ensure that judges possess final veto power. At the same, so, use of AI in the judicial system also gives rise to serious ethical and legal concerns. Issues related to the transparency and explicability of AI algorithms and their potential to entrench or magnify the biasedness built into criminal justice

[22] Year-"End Report on the Federal Judiciary, 2023 <https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf> accessed 28 March 2024".

system. Thus, use of AI in court should be regulated by well-defined ethical principles and ensured through comprehensive testing and checks.

Moreover, as AI continues to evolve and expand, the ongoing dialogue and interaction between judges, attorneys, and legislators and AI professionals, software engineers and developers, and experts in ethics, morality and human rights will be as crucial. In conclusion, the IT experts will continue working to develop the guidelines to make using AI in the courtroom possible on the most responsible and ethical levels. Such a guideline will be helpful for improving the fairness, accessibility, and performance of our court system while safeguarding the rights of all citizens.

AI-Driven Court Processes and Administrative Tasks

The maximum accuracy and efficiency of the administrative and – in the future – court procedures can be realized due to AI implementation. Making the legal system faster and more efficient is possible with the help of predictive technologies that will automate the process and decisions. It is possible to use these solutions for granting more with access to justice and simplifying the work of lawyers and court staff. This perspective allows a revolution of the court proceeding, judgment, and law research. However, even the complete accuracy and absence of emotions cannot eliminate the necessity for human, ethical issues, and sense-making in critical situations.

However, AI integration within the judicial system brings several important problems, such as bias, privacy, and accountability. Indeed, AI systems have the ability to copy and even enhance the bias present in the data they are trained on. In some cases, this leads to unfair decision and verdicts. The privacy issue arises from the use of personal data in AI oriented systems. Lastly, certain outcomes dictated by artificial means are hard to justify in some cases.

All the above clearly shows that strong ethical frameworks and guidelines should be established to regulate the use of AI in the judiciary. Besides, it includes algorithmic fairness and the fact that people's privacy should be protected and humans should exercise liability and oversight. Transparency is the key priority; the AI systems themselves should be created in such a way that it is easy to understand and run its decision process to review. Despite the fact the challenges are strong, the benefits are also very this way. AI allows one to automate labor-intensive work and helps them in making decisions; therefore, the AI redistributes resources, allowing people to focus on critical and delicate aspects of this process. Therefore, the legal profession allows offering more access to justice to handle the cases.

Chapter 4: Global Perspectives and Legal Frameworks

The deployment of the AI in investigations and the issuing of rulings is conducted differently in the USA, EU, India because of the targeted policies and rules defined for the development of the technology and its integration in the discussed spheres. These differences arise from the peculiarities of the local law and rules, the state of the market, the traditional peculiarities of the countries and the possibilities available which can influence the deployment of the AI and results of the investigations and the rulings themselves.

In the US, the adoption of AI in the judiciary and investigations is influenced by regulations and some policies. Hence, this is a critical factor that determines how the technology is being implemented and used in the sector. In the case of the European Union, as well as the EU, it plays an essential role since the respect of fundamental rights and translation of the values to the judiciary sector and investigations is also a determining factor. Finally, in India, the initiative to introduce AI into investigations and the judiciary is influenced by regulations and different cultural traditions. Thus, the difference is determined by the policies and rules that govern the use of AI in the selected countries and determining the usage cases and adoption levels of the technology in investigations and the judiciary. Therefore, it is important to know such differences to ensure that the AI technology will be applied to these critical areas responsibly and ethically.

Comparative Analysis of AI Adoption

The adoption of AI in legal procedures is gaining momentum globally, with different countries like the United States, European Union, and India exhibiting varying approaches to its integration in the judiciary and investigations[23].

United States:

In the US, AI-powered tools like COMPAS are utilized in the judiciary for risk assessment, analyzing factors such as criminal history, social background, and mental health to predict recidivism likelihood[24]. US Sentencing Commission uses AI to create and enforce sentencing guidelines for fair and just punishment. Additionally, US courts employ chatbots to provide information on court procedures, schedules, and related topics to the public, enhancing accessibility and reducing the workload of court staff[25].

European Union:

In the EU, the Smart Court system in China aids judges by analyzing past cases, suggesting applicable laws, precedents, and recommending sentences based on similar cases for informed decision-making.[26]

[23] Dovilė Barysė & Roee Sarel, 'Algorithms in the Court: Does it Matter Which Part of the Judicial Decision-Making is Automated?' (2024) 32(1) SpringerLink 117

[24] "Aditi Prabhu, Artificial intelligence in the context of the Indian legal profession and judicial system' (2023) <https://www.barandbench.com/columns/artificial-intelligence-in-context-of-legal-profession-and-indian-judicial-system> accessed" 17 May 2024

[25] ibid

[26] ibid

India:

The SC has used AI-controlled tools for processing information to help judges make decisions but not to be involved in the decision-making process. The Artificial Intelligence powered SUVAS is used in the Supreme Court of India to translate legal papers from English to vernacular languages and vice versa. In India, a more recent development is the use of AI systems in sensitive judicial decisions. An application that was cited in the judgment was "ChatGPT" which was utilized as case law with regards to bail order passed by Hon'ble High Court Judge, Justice Anoop Chitkara of Punjab & Haryana in Jaswinder Singh v. State of Punjab. [27]

Global Perspective: Internationally, policymakers are having conversations concerning the legal framework to govern the use of AI. This is because while the technology has a myriad of benefits across sectors, it is equally challenged by ethical issues. Some of the ethical and legal complications associated with AI technology include a breach of privacy, ethical dilemmas, bias in AI systems, discrimination, and security threats, to mention a few. [28] One of the ethical considerations revolves around the fact that the legal sector is grounded on human rights and values; hence, there has to be a careful limit in the implementation to assure these principles. The United States legal sector is more focused on the application of risk assessment and sentencing guidelines; the EU is more about AI in decision-making assistance, while India has made strides with several AI-powered tools in the judicial processes. This implicitly shows the differences across the world in terms of the frameworks and policies that govern the operation of AI in investigations and the judiciary.

[27] ibid

[28] "Krishna Ravishankar & Parul Anand, 'AI Judges: The Question of AI's Role in Indian Judicial Decision-Making' (2024) CALJ <https://www.calj.in/post/ai-judges-the-question-of-ai-s-role-in-indian-judicial-decision-making">

Ongoing Debates and Challenges

Data Privacy: Debates continue on balancing AI's benefits with data privacy concerns, particularly involving the acquisition and processing of sensitive information in investigations and legal decision-making.[29] Transparency in AI algorithms and decision-making processes is still a concern, with requests for explainable AI to promote accountability and eliminate biases in legal applications.[30].

Algorithmic bias: Therefore, it is vital to fight algorithmic bias in AI systems because they may enshrine prejudices of inquiries and court rulings. This issue is still a work in progress but steps towards justice and the use of AI in the legal field are implemented. These solutions range widely from sticking to the existing legislations to passing new bills and using AI is law enforcement. Much of conformance to the law and justice will depend on addressing privacy concerns, transparency, and bias issues.

Chapter 5: Ethical Considerations and Societal Impact

AI gives rise to ethical questions and has societal implications, all of which Anyinam et al. argue must be taken into account in relation to investigation and the judiciary. These issues and the societal consequences of AI on investigation and the judiciary must be adequately dealt with to ensure that the benefits of AI are not at the expense of individual rights and a just legal process. Anyinam et al. emphasize that to achieve the full potential of AI, it

[29] "Vargas-Murillo, A.R., Pari-Bedoya, I.N.M.A., Turriate-Guzman, A.M., Delgado-Chávez, C.A., and Sanchez-Paucar, F. (2024) 'Transforming Justice: Implications of Artificial Intelligence in Legal Systems', Academic Journal of Interdisciplinary Studies, 13(2), pp. 433. Available at: <https://www.researchgate.net/publication/378739404_Transforming_Justice_Implications_of_Artificial_Intelligence_in_Legal_Systems"> (Accessed: Mar 28 2024).

[30] "American Association for the Advancement of Science (AAAS), 'Artificial Intelligence and the Courts: Materials for Judges' (American Association for the Advancement of Science, 2022) <https://www.aaas.org/sites/default/files/2022-09/Paper%201_AI%20Foundational%20Issues_NIST_FINAL.pdf?adobe_mc=MCMID%3D207916409176304018114169920623059100064%7CMCORGID%3D242B6472541199F70A4C98A6%2540AdobeOrg%7CTS%3D1711586361> accessed" 28 March 2024.

must work to combat I am positive bias in AI algorithms and improve responsibility and transparency within the AI-driven legal process.

Ethical Implications of AI in Investigations

Profiling: AI-powered profiling may result in the targeting of particular communities, which may violate people's right to privacy and civil liberties.AI systems used for profiling can perpetuate biases and lead to discrimination against certain groups. There are also concerns about the accuracy and fairness of AI-based profiling.

Artificial intelligence profiling can also validate a stereotype and promote a more negative image among actual minority communities. Algorithms used by predictive policing specially target African-American and cathecumbri have caused racial discrimination to ruin various communities. Algorithmic racial profiling eliminates faith in the police and government and exacerbates minority oppression. Using racial AI profiling affirms privacy while producing an illusion of equality protected by authorities and approving arbitrary seizure and bias. Furthermore, several AI profiling methodologies relied on available data known to be false, unjust, or misrepresentation. When matched with AI technology, the data develops ableism layers that regularly ignore the individual or organization in issue. Because anything is known to the foundation is invalid, this leads to correlations and possible cruel activities.

Surveillance: The threat of information abuse and the prevention of regular people is likely to be two types of privacy shadow AI mass surveillance could generate. AI’s could permit mass monitoring and other ways to track down practically every person in the world. Furthermore, it is probable that this data could be leaked, stolen, or jumbled by malevolent hackers or other terrible characters. Surveillance is also problematic from a privacy and due process standpoint.

However, having such enormous amounts of personal data centralized in a few powerful hands, including governments and tech companies, poses the risk of abuse. AI-supported surveillance may surveil citizens to overpower their free speech, dissent, or political opposition. The course of history repeatedly demonstrated that surveillance powers quickly grew. Specifically, in times of political crises and turmoil, these powers might be severely overused.

Moreover, AI surveillance causes a chilling effect on behavior. In such a case, people change their decisions and behaviors or refuse to say something due to the threat to their privacy. Thus, it may limit freedom of expression, polarize people, and decrease trust. As a solution, there should be an optimal balance between safety and privacy, and AI surveillance should operate openly, transparently, accountable, and respecting human rights.

Potential Biases in AI Algorithms

The algorithms of AI are also discriminatory. Indeed, one possible use case of AI and machine learning AI/ML technology is to predict the future behavior of people in specific scenarios. However, not all historical data may create a comprehensive picture, leading an AI/ML model to guess or fall to some extreme interpretation, an “undue” prejudice against people’s actual distribution based on race, male, female, wealth, income, or any other relevant characteristic. Bias arises for a reason, including the data used. If the training data is historical, it also reflects a prejudice that is now institutionalized in society – then it is logical that AI learns and replicates it too. This may also give rise to risk assessment systems, which discriminate against individuals based on the demographic category rather than how dangerous they actually are to recidivate.

AI developers’ decisions can impact on what the AI learns. For instance, if an AI’s algorithm is tuned to prioritize efficiency and cost over fairness and equally, the AI might learn and reinforce existing discrimination and disadvantages against some groups. Moreover, if developers overlook how a particular consequence is being advertised, or if they happen to a homogenize team, the AI might under-advertise the said consequences. Biased AI algorithms can also be highly destructive when applied in any pattern of the justice system. For example, AI algorithms have been directly in causing longer sentences, bail settings high enough on bail to individuals from minority groups, and rejection of deportation, adding to existing inequities that defy the fundamental rights to equal treatment and proper legal procedures. Such high-stakes sectors must be subject to rigorous technical scrutiny for bias and reliant on fairness and responsibly due to the AI system used.

Societal Impact of AI on Law Enforcement and Judiciary

Law enforcement's Role: Using AI makes it possible to switch from reactive to proactive methods. With such an approach, the level of crime will decrease, and the ratio of public safety will increase. In particular, police officers can “connect the dots” by using AI-powered tools to determine patterns, resulting in more appropriate placements of police officers and early intervention to prevent crimes. For instance, predictive policing enters the data on past crimes and certain common characteristics of the associated population into an algorithm. As a result, risk maps are developed in order to indicate future criminal studies. This information will arrange community policing and collaboration between agencies and officers.

Additionally, AI tools can aid police officers in finding criminals faster and making communities safer. Although the concept of facial recognition is controversial, if used carefully and securely, it will help law enforcement identify offenders and find missing citizens. Video analytics can track persons in real time, and officers will be alerted if something awry is observed. It also assists in preventing captivity or catching it as promptly as feasible. Finally, AI-enabled tools enable policemen to focus on more critical matters.

Nonetheless, AI's use in law enforcement poses privacy, civil liberties and abuse risk. Therefore, bouncing public security against an individual must install law enforcement policies and oversight provisions that require transparent, accountable and ethical use of AI technologies. The clear guidelines on the collection, storage, or use of data should be established, safeguarding against abuse or illegal access. Regular audits and assessments check for bias and unintended consequence must take place. AI, therefore, can be a powerful benefit in supporting democratic principles and improving law enforcement's efficiency by balancing individual rights with public safety.

The role of the judiciary: There are concerns that AI-driven decision-making is fraught with biased outcomes and lack of human control, but at the same time, similar technology can enhance the performance, efficacy, and accuracy of the judicial system. Technological breakthroughs have taken place recently such that AI may even support judges and juries by enabling them to process massive amounts of information and evidence faster than before – reducing the time and resources trials and appeals consume. The technology can be used to identify relevant precedents in contrast to a case, analyze jurisprudence, formulate synopses of the essential points of contention between the parties, and correlate it with evidence. This way, the quality of judicial judgments will improve, and the mere quantity of pending cases on overloaded courts.[31]

Moreover, AI may increase the accuracy of specific judicial functions, e.g., risk and suggesting suitable penalties. Risk assessment and other AI

[31] UNESCO, AI and the Rule of Law: Capacity Building for Judicial Systems <https://www.unesco.org/en/artificial-intelligence/rule-law/mooc-judges> accessed 28 March 2024

applications can give judges a more robust and unbiased understanding of pertinent factors like criminal history, personal context and chance of recidivism. This can help guarantee that sentencing is deserved and individualized in relation to a criminal's specific threat and need, possibly decreasing the likelihood of repeated punishment.

However, there is a flip side to this medal, as with increased reliance on AI, there is less room for human judgment, and the decisions made based on AI-driven speeds are more prone to bias and errors. In particular, for AI to be applied to judicial decision-making appropriately, it must be transparent, explainable, and reviewed by humans. Judges and juries must always have the last say in making decisions, and should use AI as an additional tool to assist in this process, rather than rely solely on its algorithms and results. Assessments and audits should be regularly conducted through the entire AI-based decision-making process, to find biases prevalent and unintended outcomes. Done correctly, AI can contribute to a fairer judiciary and deliver justice more effectively while still maintaining the core principles of due process and fair justice.

Chapter 6: The Future of AI in Law and Justice

Given such potential, AI's use in law and justice may offer innovative ways to support decision-making, automate legal work, and/or engage in negotiations and adjudication. This will only be possible if four key issues are taken into account in the development of AI: human control and responsibility; ethical responsibilities and practical directions; international efforts; certain standards need to be followed. This means that the legal system can take full advantage of AI without impacting its principles of fairness and justice by focusing on them.

Potential Future Advancements in AI Technology

Enhanced Decision-Making: In the future, I see predictive analytics and sophisticated algorithms that will make the decision-making processes in the courts and during investigations in general more efficient and accurate, without being biased. For example, AI-powered tools may be employed to spot patterns and trends in large datasets, which include but are not limited to, case law, witness testimonies, and forensic evidence points. Then, between these datasets, AI generates insights and even predictions based on the combination of data points that people may see as less connected. Thus, AI may expose potential weaknesses in a case or point out the inconsistencies between witness testimony foresting or other sources, enabling investigators and prosecutors to eliminate these discrepancies. [32]

In addition, AI may be used to create risk assessment tools that can provide judges with critical information when making decisions about bail, sentencing, or parole. This kind of tool looks into several factors, including but not limited to people's criminal history and record, personal situation, and likelihood of recidivism. Hence, it can help judge whether the person is fit to live free, based on the scientific data, as opposed to guessing or stereotypes, which may result in a much more fair outcome.

Nevertheless, the use of AI for decision-making is associated with concerns over transparency and accountability. The algorithms used in those processes should be transparent and explainable, used under human oversight and review. Judges or, in some societies, juries should remain the highest authority in any controversial decision-making process. AI must be perceived as a supportive tool, but not an alternative to human reasoning or judgment. The audits and impact assessments must be carried out on a regular basis to identify the biases and side effects of AI to the decision-making processes.

[32] Michele Taruffo, 'Judicial Decisions and Artificial Intelligence' (1998) 6 Artificial Intelligence and Law 311

Automation of Legal Duties: At some time in the future, AI technology would most likely get to the level where it can complete ordinary legal duties, document inspection, and contract analysis. However, this would free up legal specialists' time to focus on even more vital and challenging aspects of the work. AI-powered systems can enable legal personnel to accomplish regular duties like document inspection and contract analysis by promptly identifying pertinent material and extracting fundamental data points. These tools could significantly cut the time and cost spent on these chores, from workers to complete more obdurate and tactical aspects of the labor.

Further, AI can be used to build smart legal research tools that use a user's query to find relevant case law, statutes, and scholarly articles. Such tools, powered by natural language processing and machine learning, would give legal practitioners access to more comprehensive and precise search results that would allow them to spend less time and effort on legal research.

On the other hand, automating the execution of these duties may result in unintended consequences, such as a problematic percentage of job loss and an increased probability of errors or misuse. As AI and machine learning revolutionize the industry, it's critical that our attorneys remain knowledgeable about these fundamental principles and constraints. As a result, they should be appropriately trained to utilize such tools and to interpret the outcome using them. The information should never be misused. captiously, authority gain access to or obtain secret legislation data.

AI in Dispute Resolution: In future, AI may become more common in the context of dispute resolution, where it might change the conflict resolution process of both mediation and the judicial system. For instance, recommendations for the conscientious use AI-powered tools case be received and test functions can be related to analyzing extensive data and evidence affiliated with a dispute. Thus, vital issues can be identified, probable resolutions offered and, with the support of machine learning-based calculations, connected with information on the likelihood of a particular outcome. As a result, participants receive a tool to make educated decisions about the benefits of approving or rejecting a settlement offer.

Another potential use of AI is implementation of online dispute resolution. Online dispute resolution is defined as a kind of dispute resolution that takes place through the use of technology. By combining negotiation, mediation, and arbitration, ODR enables the conflicting parties to come to an agreement remotely. The platforms can also use AI to carry out certain aspects of dispute resolution, such as document exchange and communication, and scheduling. This will make ODR more accessible and efficient, especially for small-scale disputes not suitable for traditional litigation.

Conversely, the past of AI used in dispute resolution involves unfairness, lack of transparency, and growing bias. It is vital that AI-based disputes resolution methods build comprehensive protection tools for the parties involved. Such measures need the algorithm to be transparent, explainable, and controlled manually through human behavior and adjusted to regulations. Regular impact assessments and audits should be conducted to identify and remove bias and unintended consequences of AI's implementation in the resolution process.

Human Oversight: However, in order to fight potential fraud, it is important to retain human oversight in AI-driven legal processes to make certain that it is accountable, transparent, and non-discriminatory in the decision-making process. Legal professionals and policymakers need to make sure and ensure human oversight, which includes human accountability, will remain a part of AI-driven legal processes. This means that the final decision still lies with the judges and juries, as AI can just be a support tool and should not replace human intelligence, control, and decision-making. Regular audits and impact assessments can help to control and eliminate possible AI bias and ensure no harm is done in legal processes.

Legal practitioners should have a thorough understanding of the fundamentals of legal rules and the constraints within which AI tools work. Training and funding for training and research on the application and results for AI instruments should be given to all legal practitioners to stop the use of AI in litigation and ensure it is always under human direction.

Ethical Aspects: Eliminating prejudice and protecting people's privacy and the right to a fair trial all need to be central. The ethical principles of development and use should largely be used to formulate ethical AI guiding rules in the legal environment. This means creating AI systems that will limit the occurrence of prejudice and discrimination as well as which will enhance individual privacy and promote the withholding of fundamental legal principles on apt punishment and such. Regular audits as well as impact evaluations are assessments that should help identify and mitigate ethical problems.

Furthermore, legal professionals and policymakers should strive to create transparent standards and rules of AI use in the legal context. Similarly, these should establish requirements for the necessity and adequacy of data and protocols of AI system development and deployment. If legal professionals promoted integrating rules of AI use from the ethical perspective, the beneficial potential of AI would be realized while preventing or minimizing all its probable risks and harms.

International Cooperation and Standardization for the Development of AI

Globally, collaboration primarily entails the development of similar standards and policies on the use and development of AI in the legal process to guarantee the same conformity, inter-operability, and morality. In building similar standards and policies on the use of AI in legal processes, international cooperation is needed in every country. All countries cooperate to acquire a set model that determines the terms and conditions that guide the execution and development of AI systems responsibly. This is used to secure the artificial entity and enables it to be deployed and used morally and consistently in every jurisdiction to assure the public on its trust.

International cooperation is more critical than ever to share the knowledge, resources, and best practices in AI for legal processing. When the experience of common errors and victories is at hand, it is less complicated for countries to create the ideal AI systems in terms of being beneficial for the citizens. The probability of progress and innovation is also higher as the countries can unite to create more sophisticated AIs that could make unrealistic tasks for a single country possible.

Standardization: Standardization could also guarantee responsible and trustworthy use of this technology. One way to use standardization to foster a belief and responsibility in this technology is to develop guidelines that that govern court and investigation applications. For instance, standardized drive, standards for cot deployment, and disinterested data share can foster a belief toward AI as a reliable system and proper for usage. Standardized guidelines can be applied to ensure responsible and efficient AI uses. This method addresses issues regarding algorithmic biases and data privacy and transparency. Which can help in allaying everyone’s concern and build trust and confidence in the fairness and justice of AI-redressed judicial measures.

Additionally, Standardization is also used to encourage interoperability and compatibility. This goal is achieved by encouraging the use of a single set of technical standards and protocols that all AI is based on, some of which support data and information sharing in a seamless manner. Consequently, it is possible to share case and investigation information from different jurisdictions without difficulty. Therefore, fighting crime and trading justice through international cooperation and collaboration could be made easier.

Chapter 7: Conclusion

The significance of adopting a human-centered approach.

AI's advantages include enhancing productivity and accuracy, or substantially simplifying inquiry as well as judicial action. However, there is a continuing need for ethical and responsible AI integration due to privacy issues, bias in algorithms, and insufficient transparency. Qualified and informed human oversight is essential to ensure accountability, transparency, and fairness in the decision-making process for AI-based legal procedures. Ethical concerns should be primary and demanding throughout AI creation and application to prevent bias, protect privacy rights, and guarantee even-handed adjudication in cases.

Standardization and global collaboration

Another important necessity is that the countries form and develop relationships. In this way, consistency, interoperability, and high ethics of the justice system are well assured. This involves properly published standards and guidelines on how AI should be developed and used. Asustek et al. states that informal protocols can terminate the issue of data privacy and prevent new cases of algorithmic bias and lack of transparency. Therefore, the software can be trusted, and the society trusts it.

AI affects in two ways the field of investigation and the judiciary. The first one concerns the pros and cons of efficiency and effectiveness. The next issue is associated with the pitfalls of algorithm bias and the lack of transparency concerning data privacy. The answering question on the ethical aspect of using this technology to attain justice refers to the human-centered approach that takes humans’ moral judgment and supervision as a default. Additionally, international cooperation is essential in securing a level playfield around these AI implementation tactics to ensure adequacy in addressing the issues of credibility in legal processes.

GOVERNANCE OF ARTIFICIAL INTELLIGENCE IN CRIMINAL JUSTICE SYSTEM: A LEGAL AND ETHICAL ANALYSIS

Dr Anisa Shaikh and Ms Vipasha Chirmulay[1]

Abstract:
The courts in India have been encouraging the adoption of Artificial Intelligence ('AI') in the legal field with a measured and reasonable stand. The Punjab & Haryana High Court has acknowledged the use of ChatGPT for legal research, with the expectation that it can enhance efficiency. Former Chief Justice of India, D.Y. Chandrachud has emphasized that AI will supplement rather than replace traditional legal skills. As AI technologies become increasingly integrated, the need for robust and ethical governance frameworks becomes important to ensure fairness, accountability, transparency, and the protection of fundamental rights.

To analyze the emerging role of AI in criminal justice systems globally and in India, this paper focuses on the legal principles and policy proposed for its governance. The paper aims to critically examine the inclusion of AI in the criminal justice system supported by Indian judiciary within the criminal justice domain. It may serve as a reference work for future legal research, policy discussions, and academic curriculum development in law and technology fields. The paper will advocate for a governance model that prioritizes human rights, promotes public trust, and fosters continuous evaluation and adaptation.

Keywords: criminal justice, India, law, artificial intelligence, courts

[1] Assistant Professors, New Law College Pune.

"We need to deploy the technology in ways that enhance rather than erode public confidence in the justice system."

- Marc A. Levin and Jesse Rothman

Introduction

The integration of Artificial Intelligence (AI) into the criminal justice system marks a transformative era, promising enhanced efficiency, accuracy, and objectivity across various critical functions. From predictive policing and sophisticated forensic analysis to risk assessment for bail and sentencing, and even correctional management, AI-powered tools are increasingly being deployed worldwide. In India, the adoption of AI-based systems like the Crime and Criminal Tracking Network & Systems (CCTNS), Inter-operable Criminal Justice System (ICJS), and the Supreme Court Portal for Assistance in Court's Efficiency (SUPACE) is rapidly expanding, aiming to address the case backlogs (over 50 million pending cases). Globally, the predictive policing market is projected to reach 5.6 billion dollars in 2025 and 196.3 billion dollars by 2034, which points to the investment and reliance on these technologies.2 This pervasive integration offers benefits in tackling modern crime challenges and improving public safety, but also leads to ethical, legal, and societal implications.

AI tools can analyze huge amounts of data to identify patterns and trends, enabling law enforcement agencies to predict and prevent crimes more effectively. Real-time data analysis can assist with recreating crime scenarios, identifying wanted individuals, and mitigating new threats. They can automate routine legal tasks, such as legal research, document review, and case management, thereby speeding up judicial processes and reducing backlog in courts. AI algorithms can assist judges, prosecutors, and defense

[2] Dimension Market Research, *AI in Predictive Policing Market* (2024), https://dimensionmarketresearch.com/report/ai-in-predictive-policing-market/ (last visited July 28, 2025).

attorneys in evaluating evidence, assessing risks, and predicting case outcomes based on data analytics. AI technologies, such as facial recognition, predictive analytics, and surveillance systems can ensure public safety and national security.

The proliferation of AI in this domain is not without its challenges. Recent trends highlight growing concerns regarding algorithmic bias, the 'black box' problem of opaque decision-making, and the potential for these systems to perpetuate existing societal inequalities. For example, studies in the United States have indicated that AI tools used in sentencing have, despite promises of objectivity, appeared to discriminate against certain demographic groups. In India, where societal biases related to caste, religion, and socioeconomic status can influence judicial outcomes, the uncritical implementation of AI in policing and other areas poses a significant risk of entrenching rather than mitigating these biases. The absence of comprehensive, AI-specific regulatory frameworks in many jurisdictions, including India where the legal landscape is still evolving, further exacerbates these concerns. This paper has critically analyzed the governance challenges associated with AI in the criminal justice system, and the balance between technological advancement and the fundamental principles of fairness, accountability, transparency, and human rights.

Ethical and Societal Challenges

AI promises enhanced efficiency, accuracy, and objectivity in various facets of law enforcement, adjudication, and corrections, however its deployment raises concerns regarding fairness, accountability, transparency, and the potential for exacerbating existing societal inequalities. This section discusses issues such as algorithmic bias, privacy infringements, the erosion of human discretion, and the implications for human rights and social justice in an AI-driven criminal justice system.

Algorithmic Bias and Discrimination

Algorithmic bias takes place when systematic errors in machine learning algorithms produce unfair or discriminatory outcomes. When formulas include prior arrests or legal history, past discrimination can be repeated with the algorithm, which will influence who the police target. When police encounter a high threat score, this increases the rates in which they use force. Databases can also have limitations, which lead to the misidentification of people in certain groups. If the data is used in facial recognition technology, it can misidentify suspects and incarcerate innocent people.3

Lack of Transparency and Explainability

AI algorithms, particularly deep learning models, process vast amounts of data and produce outputs, but their inner workings are too complex to understand. Governments have constitutional obligations to be transparent, in upholding citizens' rights to due process and equal protection under the law. Accountability ensures that there are clear mechanisms to assign responsibility and provide redress if these systems cause harm. Risk assessment algorithms that inform judicial decisionmaking in sentencing should be built and operated on an open source software platform.4 AI algorithms may not be accurate which can lead to unfair decisions or wrongful convictions. Increasing reliance by judges, lawyers, and defendants could reduce their ability to exercise discretion and make judgements.5

[3] Ave Maria L. Sch., *Can Algorithms Lessen?,* Ave Maria L. (Apr. 22, 2024), https://www.avemarialaw.edu/can-algorithms-lessen/.

[4] Caleb Watney, *When It Comes to Criminal Justice AI, We Need Transparency and Accountability*, R Street Inst. (Dec. 1, 2017), https://www.rstreet.org/commentary/when-it-comes-to-criminal-justice-ai-we-need-transparency-and-accountability/.

[5] Ivas Konini, Rokaj, *The Challenges on Implementing Artificial Intelligence in the International Criminal Justice System* (May 2024), https://www.researchgate.net/publication/380360976_The_Challenges_on_Implementing_Artificial_Intelligence_in_the_International_Criminal_Justice_System.

Difficulty in understanding and challenging decisions made by AI systems violates the principles of due process and fair trial.6

Accountability and Liability

When an AI system makes a mistake that leads to a wrongful arrest, an unfair sentence, or other harmful outcomes, determining who is legally responsible is incredibly complex. AI systems cannot be held liable in the same way as humans. For damage caused by a high risk AI systems and the need to compensate the injured person in civil law, there are discussions on attributing legal personhood or capability to AI systems for the purpose of holding them accountable. The stakeholders include the persons that design, program, and monitor the AI system, the users as well as those interacting with the system. There should be different expectations towards each person involved in the design, training, production, or use of an AI system depending on their tasks.7 Concepts like mens rea (guilty mind) and actus reus (guilty act) become problematic when an AI is involved, because AI does not possess consciousness or intent.

Privacy Concerns

AI systems may require extensive data collection, which raises issues about privacy and confidentiality of defendants in trials. Facial recognition technology may result in mass surveillance, and can infringe on individual privacy. Cybersecurity threats can pose a threat to confidential information such as criminal records, resulting in misuse, identity theft and fraud. Predictive policing activities can lead to unethical profiling and surveillance

[6] Ivas Konini, Rokaj, *The Challenges on Implementing Artificial Intelligence in the International Criminal Justice System* (May 2024), https://www.researchgate.net/publication/380360976_The_Challenges_on_Implementing_Artificial_Intelligence_in_the_International_Criminal_Justice_System

[7] Athina Sachoulidou, *AI Systems and Criminal Liability: A Call for Action* 11 Oslo Law Review 1 (2024), https://www.scup.com/doi/10.18261/olr.11.1.3

based on age, race and gender of individuals in a community. Biometric or online activity data can be intrusive surveillance.8

Emerging Governance Framework

The integration of AI within the criminal justice system necessitates the development of robust and adaptable governance frameworks. While AI promises enhanced efficiency, accuracy, and potentially fairer outcomes through applications like predictive policing, risk assessment, and automated legal research, its deployment also introduces ethical, legal, and societal challenges.

Global framework

The Council of Europe Framework Convention aims to ensure that AI systems, regardless of whether they are used by public authorities or private actors, align with human rights, democracy, and the rule of law. The European Union's Artificial Intelligence Act introduces a tiered system, categorizing AI systems based on their risk levels (unacceptable, high, limited, minimal) so that each has different requirements. Legal frameworks need to address these issues through algorithmic transparency, bias detection, and mitigation.

Courts have been consistently acknowledging that the use of AI in criminal justice is transparent and its decision-making processes can be scrutinized:

Procedural Fairness: The right to know and challenge the data, reasoning, and output of any AI system used in judicial proceedings is a recurring judicial theme.

[8] Ivas Konini, Rokaj, *The Challenges on Implementing Artificial Intelligence in the International Criminal Justice System* (May 2024), https://www.researchgate.net/publication/380360976_The_Challenges_on_Implementing_Artificial_Intelligence_in_the_International_Criminal_Justice_System

Bias and Accountability: Courts stress the risk of algorithmic biases and the need for independent audits and safeguards against unfair discrimination in AI-driven decisions.

The intersection of artificial intelligence and criminal justice governance has come under increasing judicial scrutiny as discussed in the following cases:

1. State v. Loomis (Wisconsin, USA), Wisconsin Supreme Court, 2016:

This is a case regarding the use of the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm in sentencing decisions. The defendant, Eric Loomis, challenged the use of COMPAS on grounds that it was opaque, non-transparent, and did not allow defendants to contest the algorithmic assessment, raising due process concerns. The court upheld the use of COMPAS but required that judges be made aware of its limitations, especially regarding gender and racial fairness. Loomis is widely cited as establishing both the opportunities and challenges in utilizing AI-driven risk assessment tools in judicial processes, specifically highlighting the need for transparency and algorithmic accountability.

2. Crawford v. Washington (USA), US Supreme Court, 2004:

Although not directly about AI, this case is often referenced for the principle of the right to confront evidence, an issue raised in the context of AI-generated evidence or risk scores. It amplified concerns about defendants' rights to contest and understand evidence produced or processed by algorithmic tools. It underscores the procedural due process concerns when AI tools are used, especially regarding the "black box" nature of evidence generation.

3. Townsend v. Burke (USA), 1948:

This case established the principle that the accused must be informed of the basis for their conviction or sentence. This due process right is invoked in contemporary debates over opaque AI systems in sentencing, where the rationale cannot be fully disclosed or explained to the accused. It forms the foundation for arguments insisting on transparency and explainability in AI-driven court decisions

.

4. Puloka and Arteaga (EU):

Cited within European scholarship and legal commentary as examples addressing constitutional issues surrounding AI in criminal justice, including due process, equal protection, and privacy, courts here have increasingly required that AI uses in judicial decision-making comply with human rights and allow for contestability, often referencing the European Convention on Human Rights.

Common principles

Common principles of governance are important to build trust and prevent harm, especially when AI is used for tasks like crime prediction, evidence analysis, and sentencing. Microsoft has identified six principles to guide the development and use of AI such as fairness, reliability, privacy, inclusiveness, transparency, accountability. In 2021, the Niti Ayog released an approach document on Principles of Responsible AI.

Transparency: AI systems and their decision-making processes should be understandable to stakeholders, allowing for scrutiny and accountability. The Supreme Court of India has held that transparency in decision making is critical even for private institutions, as the Constitution guarantees accountability of all State action to individuals and groups. The design and functioning of the AI system should be recorded and made available for external scrutiny and audit to ensure the deployment is fair, honest, impartial and guarantees accountability.[9]

Fairness: AI algorithms should be designed and trained to avoid perpetuating existing biases in data, ensuring equitable outcomes. The development and use of AI systems must consider both substantive and procedural fairness by giving lawful reasons or justifications for their output.[10]

Accountability: Mechanisms should be in place to identify and address errors or misuse of AI systems, ensuring someone is responsible for their actions. If there are no consequences, there will be no responsible action. An AI system which has multiple roles behind individual decisions makes it difficult to attribute errors, find the cause of action and assign liabilities. Stakeholders should conduct risk and impact assessments to find out the direct and indirect impact of AI systems, and set up an auditing process.[11]

Privacy: Data collection and usage by AI systems must comply with privacy regulations and protect individuals' personal information. Technology can record and analyse an individual's personal life without their consent or knowledge.

India's legal responses to AI in the criminal justice system

India's approach to regulating AI in the criminal justice system is still evolving, to adapt existing laws and new policies to address challenges and opportunities presented by AI. The Information Technology Act 2000 addresses data protection, cybersecurity, and cybercrime, which are indirectly

[9] NITI Aayog, *Approach Document for India Part 1 – Principles for Responsible AI* (2021), https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.
[10] NITI Aayog, *Approach Document for India Part 1 – Principles for Responsible AI* (2021), https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.
[11] NITI Aayog, *Approach Document for India Part 1 – Principles for Responsible AI* (2021), https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.

relevant to AI applications in the criminal justice system. The Supreme Court's ruling on privacy as a fundamental right under Article 21 of the Constitution has implications for AI applications, particularly in surveillance and data collection. It has held that the right to privacy is an intrinsic part of the right to life and liberty guaranteed under Article 21.

The Digital Personal Data Protection Act 2023 is the primary law regulating data collection, storage, and processing. It does not address algorithmic biases or AI-generated data misuse, nor AI audit and accountability. In a 2023 bail order before the Punjab and Haryana High Court, Justice Anoop Chitkara referenced OpenAI's ChatGPT, leading to debates about judicial reliance on AI for sensitive decisions. This is recognized as a first for Indian judicial decision-making involving AI, raising questions on technological reliability, fairness, and due process under the Indian Constitution.

Recommendations and Best Practices

The increasing integration of AI within the criminal justice system presents both unprecedented opportunities for efficiency and critical challenges to fundamental rights and ethical principles. As AI tools evolve from predictive policing algorithms to forensic analysis and sentencing support, ensuring their responsible and equitable deployment becomes paramount. This section offers a set of recommendations and best practices to guide policymakers, legal professionals, and technologists. In the absence of a legal framework, AI can be used punitively and excessively leading to misuse.

Human-Centric Approach

Human centred AI is a branch of AI which seeks to create AI systems that augment, prioritize and not displace human abilities, needs, and values.This means that fundamental rights of humans must be enhanced by the use of AI in the criminal justice system, and must be protected from harm. Predictive policing systems can target minority groups and affect their security, health and fundamental rights. AI systems should protect vulnerable groups from being targeted and not reinforce discrimination. They should not impact the neutrality of the judgements by maliciously manipulating the decision making process.12

Bias Detection and Mitigation

Unidentified, incomplete, outdated, inaccurate data should be removed from the data sets. However, complex and diverse data may represent real world situations. This conflict affects social perceptions of fairness and justice. The outcomes should be monitored and corrected so as to ensure that biases do not worsen as the system learns from the new data. The target audience should be represented in terms of gender, ethnicity and other grounds. Documentation, testing and risk management are important to ensure that there is no algorithmic discrimination.13

Transparency and Explainability

There is emphasis on transparency and accountability in order to recognize how AI systems take decisions. The rationale behind understanding AI decisions is to understand the factors and data that influence the outcomes of a system. Explainability means that the AI system should be able to reason why it gave a particular outcome, whereas transparency requires the AI

[12] Deniz Çelikkaya, Mustafa Karayigit, *The Use of AI in Criminal Justice: Unpacking the EU's Human-Centric AI Strategy* 8 Nordic Journal of European Law (2025), https://www.researchgate.net/publication/390479556_The_Use_of_AI_in_Criminal_Justice_Unpacking_the_EU%27s_Human-Centric_AI_Strategy#:~:text=human%2Dcentric%20AI%20strategy%20i,following%20key%20implications%20thus%20emerge.&text=designers%2C%20algorithms%20may%20take%20an,correlations%2C%20inferences%20and%20interpretations%20made

[13] Deniz Çelikkaya, Mustafa Karayigit, *The Use of AI in Criminal Justice: Unpacking the EU's Human-Centric AI Strategy* 8 Nordic Journal of European Law (2025), https://www.researchgate.net/publication/390479556_The_Use_of_AI_in_Criminal_Justice_Unpacking_the_EU%27s_Human-Centric_AI_Strategy#:~:text=human%2Dcentric%20AI%20strategy%20i,following%20key%20implications%20thus%20emerge.&text=designers%2C%20algorithms%20may%20take%20an,correlations%2C%20inferences%20and%20interpretations%20made

system to disclose its functions. AI systems must follow a transparent model which can be audited and tested for bias during development.

Robust Data Governance

Accurate datasets should be provided to train the algorithms. Protecting the data of individuals is important to protect the security of individuals. AI data systems are vulnerable to cyberattacks and may even be used to commit cyberattacks. Clear safeguards and policies regarding AI applications in criminal justice domains must be maintained.14 AI-based data must be encrypted and stored securely to allow authorized personnel access, periodic assessments must be conducted and sensitive data must be redacted before it is released to the public.

Legal and Ethical Audits

When AI errors lead to wrongful imprisonment or unfair trials, the justice system may encounter legal challenges such as appeals, retrials, or claims for compensation. Focusing on liability and due process under AI based legal mechanisms, requires a legal framework to preserve fundamental rights with fair treatment. Criminal liability will be attributed if negligence is a component of mens rea. The burden of proof will still lie with the prosecution to establish the guilt of the accused. AI generated content may hold the risk of manipulating the evidence and questioning the authenticity of information presented in court.15

[14] Mohd. Khairul Ahmad & Mohamad Izwan Bin Ishak, *DATA PROTECTION, PRIVACY, AND SECURITY IN THE CONTEXT OF ARTIFICIAL INTELLIGENCE AND CONVENTIONAL METHODS FOR LAW ENFORCEMENT,* (Aug. 2023), available at https://www.researchgate.net/publication/373459224_DATA_PROTECTION_PRIVACY_AND_SECURITY_IN_THE_CONTEXT_OF_ARTIFICIAL_INTELLIGENCE_AND_CONVENTIONAL_METHODS_FOR_LAW_ENFORCEMENT (last visited July 28, 2025).

[15] Vartika Dixit & Niti Singh, *Artificial intelligence and criminal liability in India: exploring legal implications and challenges*, 2024 INT'L J. OF SCI. & RES. (IJSR) 1, https://www.researchgate.net/publication/379955663_Artificial_intelligence_and_criminal_liability_in_India_exploring_legal_implications_and_challenges.

Conclusion

The regulation of AI in the criminal justice system has become essential. Strong legal structures, ethical principles, and effective oversight mechanisms are crucial to leverage AI's benefits while avoiding the exacerbation of historical inequalities. Through the implementation of risk-based regulations, clear impact evaluations, and inclusive oversight, regions can promote AI advancements that improve safety and equity while protecting the fundamental rights of all individuals under state authority

The regulation of AI tools in criminal justice systems needs a systematic approach that addresses the protective fundamental rights and the technological advantages offered. There is a need for strong governance legal frameworks for AI due to exhibited risks.

Governance strategies for AI must revolve around human factors that should reinforce the idea that technology should augment but it should not substitute the human decisions. The far-reaching repercussions of bias embedded within algorithms and the clockwork-calculation in the surface of AI's decision-making require ongoing oversight, proactive auditing and significant transparency efforts.

To preserve ethical principles in AI technologies, governance frameworks must adapt simultaneously. Emphasizing the experience of early adopters reveals critical insights towards building effective and fair AI systems within the criminal justice landscape. Achieving this goal demands careful coordination between legal scholars, policy developers, technologists, and community stakeholders to guarantee that justice is served and not AI inappropriately used to subvert it.

Innovation and protection of fundamental rights are safeguarded with empirical evidence to guide policy decisions and implementation strategies. Such an AI method approach will enhance criminal justice with integrity and fairness in the legal systems.

There is a need to create statutory legal bodies to govern AI and its use in the justice system for transparent and effective application.

There is a need for proper evaluation of AI systems for protection of fundamental rights before its deployment for fair and biased impact.
A common legal framework can be created for measuring the performance of AI and its accountability.

There must be human control over AI assisted legal decisions because full automation is not possible due to the nature of law and its enforcement.
There is a need of an hour to train judicial and enforcement machinery on AI systems and its limitations.

There is a need to frame uniform global AI governance principles and standards in criminal justice systems in administration of criminal justice systems.

ARTIFICIAL INTELLIGENCE (AI) & ITS CHALLENGES FOR INTELLECTUAL PROPERTY RIGHTS (IPR)

Mrs. Mayura Pawar and Mr. Harshad Pujari[1]

Abstract

Artificial Intelligence (AI) can be defined as the intelligence exhibited by the machines. The trend of AI is becoming popular in various sectors like Education, Medical Science, telecommunication, aerospace, robotics etc. these invention related to AI generally uses deep learning and neural network as techniques for the development and Robotics is something which is becoming popular around the world and it is seen that the thing which become popular requires protection, this article also lays emphasis on the patent prosecution and protection of the robotics work. In recent times AI is facing a lot of issues like proprietary issues of Inventor ship and inadequate regulations which have raised many questions. This article focuses on the possible solutions to the issues related to AI and recent developments that took place in the copyright law with respect to Artificial Intelligence (AI). It also touches upon compatibility of the Indian copyright law in handling the work created by Artificial Intelligence (AI) and the possible issues that may arise if AI is acquired as a separate legal entity. Apart from this there are complexities in acknowledging AI as the author of Copyright. The researcher has also discussed the importance of patent in the protection of AI generated inventions and challenges that AI comes across relating

[1] *(Assistant Professors) Bharati Vidyapeeth Deemed to be University New Law College, Pune*

to the IP policy like ownership, infringement etc. In this article the researcher has reviewed various juristic challenges to AI and has given appropriate as well as practical suggestions to overcome such issues.

Keywords: Artificial Intelligence, AI & copyright, AI & patent, Robotics, IP challenges, Inventor ship.

1. Introduction

Artificial Intelligence (AI) technologies have become ubiquitous in our daily experience as consumers in this digital era. AI is more and more used by business in order to understand and influence consumer habits directly or indirectly. Society including consumers are gaining more awareness about this practice, which is increasingly looked at with concern and uneasiness, as being a source of risks for consumers, whose privacy, autonomy and wellbeing are threatened. "Artificial intelligence is likely to be either the Best or Worst thing to happen to humanity"[2] on this note in the digital era of transforming world with different technologies as well as different types of crimes are on rise still it is not very strange for the human society that their places are taken by the robots and software's for doing household works, industrial works and any form commercial works.

Among the Nations across the world, 20th century was accredited for the effect of adoption of Liberalization, Privatization and Globalization, whereas 21st century will be distinguished for impacts that the advancements in technologies across fields would create. Artificial

Rory Cellan Jones, "Stephen Hawking warns artificial intelligence could end mankind", *BBC*, 2 Dec. 2014, available at: https://www.bbc.com/news/technology-30290540 (last visited on 8 Dec. 2024).

Intelligence will be one among those advanced technologies contributing major share in it. Through its functioning in multidisciplinary sectors, artificial intelligence technologies forms part of the society by making life easier. Artificial intelligence is both compelling and controversial, not for its practical achievements, but rather for the metaphor that lies behind the programs. On one hand, this technology enables the people to do works which are hazardous to human life and to increase productivity and economies in the market. On the other, this intelligent system forces the developing countries, like India, to encounter the risk of regulating it in the practical world. As per NITI Aayog's report in 2018, India, being developing economies in the world, has now initiated its process in enabling such intelligent systems across various sectors such as health, agriculture etc. In spite its innovating ideas, regulatory framework and implementation of such system in each sector involves in-depth research and employment of Data accordingly.

Though there is no commonly accepted, rather defined definition of AI, according to Lexico, an Oxford Dictionary, Artificial Intelligence is a "The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between language"[3]. John McCarthy, who was considered to be father of Artificial Intelligence, defined the Artificial Intelligence as "the science and engineering of making intelligent machines"[4].

[3] "Meaning of Artificial Intelligence", *avaiilable at:* https://www.lexico.com/definition/artificial_intelligence.

[4] Human Rights in the Age of Artificial Intelligence, Professor John Mc Carthy's defination of AI, *available at:* https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf

According to Gartner, Big Data is data that contains greater **v**ariety arriving in increasing **v**olumes and with ever-higher **v**elocity. Here variety refers to the various sorts of data that are available, velocity is the fast rate at which data is received and acted on and volume is nothing but the amount of data[5]. Machine learning, on the other, is nothing but a part, under the umbrella of Artificial Intelligence and is related to each other. According to Arthur Samuel, Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed. Or machine learning is a branch of artificial intelligence that systematically applies algorithms to synthesize the underlying relationships among data and information[6].

2. Recent Development in Copyright With Respect To Artificial Intelligence (AI)

The upcoming generation is highly addicted to technology driven things whether it be smart phones, laptops and various other gadgets. According to statistics the total number of smart phone users in the world is 3.8 billion, which is 48.37% of the total population in the world.[7] Role of Artificial Intelligence is very important in the upcoming years for the evolution of modern technologies as well as software. If a person is a Marvel Fan he/she can relate to the AI very much. The same thing what Iron Man did in the movie the same thing is now becoming a reality.

The concept of artificial intelligence dates back to 19th century, where in the Dartmouth conference (1956) the emphasis on the artificial intelligence was laid. This concept alludes to an algorithmic instrument

[5] "The Defination of Big Data", *available at:* https://www.oracle.com/in/big-data/what-is-big-data.html

[6] Awad M. & Khanna R. "Machine Learning. In: Efficient Learning Machines", *Springer*, Apr.27, 2015 https://doi.org/10.1007/978-1-4302-5990-9_1.

[7] How-many-phones-are-in-the-world, *available at*: https://www.bankmycell.com/blog

that can figure or can replace human insight needed in the fruition of specific errands. Albeit the publicity around the popular expression "Artificial Intelligence" is thought or focused with the terms like Software, Deep Learning, Algorithms, and so on. It is substantially more than that. Computerized reasoning or artificial intelligence is not any more confined to processing calculations, robots, Alexa, and so on, in easier terms, it is presently not, at this point an apparatus or PC program which is subject to human orders and mediation. Maybe it has now been changed, to settle on an autonomous and inventive choice with no human mediation.

If we see today AI is there in every field whether it be Gaming, Music, and Movies etc. AI has brought a revolution in these field. Creators' are now using AI to generate products related the field like gaming, music and movies since AI is becoming popular, the work of the creator needs to be protected. So the question arises whether the work generated by AI is protected under copyright Act or not? Another question that comes in the mind is that who owns the authorship of the work created by AI? The main purpose or we say the main objective of the copyright law is to protect the work of the creator as author is the first owner of copyright[8]. In the Indian context the objective of the copyright law is to protect the expression of the original work and not the idea of the work. Section 2(d) (IV)[9] talks about work made by PC and further gives that the individual who is mindful to cause the work is the "creator". This arrangement some way or another deals with the possibility of AI-produced work with human mediation, and plainly assuming there is human intercession in the work created, possession remains alive with the developer. Yet, it isn't eloquent about the works made by AI without human intervention.

[8] *Section 17*, Copyright Act 1957
[9] Copyright Act, 1957

This in some way or another prompts a hazy situation that is as yet unanswered. The Hon'ble Court in Rupendra Kashyap Vs. Jiwan Publishing House Pvt. Ltd.,[10] held that with regards to address papers for an assessment, that the creator of the assessment paper is an individual who has gathered the inquiries; the individual who does this incorporating, is a characteristic individual, an individual, and not a counterfeit individual; Central Board of Secondary Education is certifiably not a characteristic individual and it is qualified for guarantee copyright in the assessment papers just on the off chance that it builds up and demonstrates that it has drawn in people explicitly for reasons for readiness of aggregation, known as question papers, with an agreement that copyright in that will vest in Central Board of Secondary Education.[11] Similarly, in different decisions the courts have seen that a juristic individual is unequipped for being the writer of any work in which copyright may exist.[12] This is likewise settled with the Practice and Procedure Manual (2018) gave by Copyright Office, which plainly expresses that with the end goal of Copyright, just regular individual subtleties should be given as Author of the work

The question arises what if the ownership is granted to the AI generated work is created without any human intervention? The question can be answered if we know how AI generates the work and how it produces it. Nowadays computer have been an incredible supporter of the creation of literary works and artistic works for seemingly forever, but witnessing the growth of IT over the past few years we must reconsider the elements of Computer and creative learning. In the present times Machine learning

[10] 1994 (28) DRJ 286

[11] *Navigators Logistics Ltd. vs Kashif Qureshi & Ors*. {CS(COMM) 735/2016}

[12] *Tech plus Media Private Ltd. vs. Jyoti Janda*, {(2014) 60 PTC 121}; *Camlin Pvt. Ltd. vs National Pencil Industries*, (AIR 1986 Delhi 444).

has become the subtype of AI, this enable the machine to do what is being fed by humans with no other programming done by themselves.

Computer programs and Machine Learning are straightforwardly relative to one another, as PC programs are made to fill the need of AI. The data or the algorithm are designed in such a way so that the machine takes the input from the data, process it, and then makes the choices or the decisions likewise. The data are provided to them by the programmers, the machines learns from these data and make the independent decisions. But in any case, this doesn't conceal the fact that AI regardless of whether proficient to settle on an autonomous choice isn't reliant on human-took care of calculations and programming, the work that is made is only a high level/altered variant of the given info took care of by the programmer. Thus it is secure to say that AI-produced work lacks originality.

The AI works even does not fall within the ambit of Doctrine of Modicum of Creativity[13] as this doctrine states that the work is original if it has sufficient creative skills and judgement involved in it. But in the case of AI we see that the machine only performs only what is being fed by the humans in the public domain, it does not have the personal judgement skills in it.

Now answering the question as to whether the AI works are given protection under the copyright law we can say that current laws and doctrines are not very well equipped in handling the AI. And apart from this India laws has given more importance on the work which involves human involvement. With regards to the Indian set of laws, it has

[13] Propounded in U.S. and it dates back to 1884

allowed legal status to numerous artificial individuals like organizations, associations, and so on however AI doesn't cut, as it is just seen as an apparatus for human help. The advancement of AI has progressed significantly yet it is as yet premature in its turn of events. This is an unknown area for the researchers of law as the making of AI is moderately new. So there is no very much characterized law administering the circle of copyright issues.

Having placed in a word for AI, there is a sure likely issue that may manifest and those are identifying with Copyright Infringement and liabilities. To decide how AI might manage the above issue first, we need to watch out if there is any conceivable way out. There can be an alternate area referenced in the Act that manages the work produced by AI explicitly.

There can likewise be a part that decides the subject of creation and possession in such manner like:

- On account of AI-produced work made without human intercession, the origin can be given to the proprietor for example the developer and the possession can be given to the actual AI.
- Additionally on account of works created by AI with human mediation the origin can remain alive with AI and the proprietorship with the developer.
-

Yet, as the well-known adage goes simpler than said to be done, to consider every one of the admirable statements the primary concern required is that the Copyright laws should perceive AI as a legitimate element. Whenever done then it will be not difficult to decide the issue identifying with Copyright Infringement and liabilities. Presently having talked about the extent of giving AI works copyright we should perceive what might be the plausible cons of giving the AI copyright.

The law in India isn't sure about this subject however for the conversation let us say speculatively that AI is acknowledged as a lawful element and the works made by it very well may be protected. A portion of the issues that might be confronted are as per the following:

- Inventorship: Originality of work is one of the necessities for copyright. One of the significant concerns concerning AI-made work is that numerous individuals accept that AI-made work isn't unique work since AI is a human-composed program and it is restricted by the boundaries entered by the software engineer.
- Issue of encroachment: AI itself is certifiably not a lawful element so on the off chance that there is an encroachment with respect to the AI it will be extremely hard to put the risk
- Moral rights: The creators has monetary right as well as moral rights also. Moral rights incorporate the privilege to respectability and the privilege to paternity. Assuming AI is made by the creator, these rights will lose their importance as these rights are intended to oblige the feelings of the creator since AI doesn't have any feeling these rights become excess.
-

According to the current law of copyright in India each copyright holder is qualified for Royalty this privilege can't be deferred. Where AI is the creator there are numerous unanswered inquiries, for example, who will set the royalty sum and how the sum might be dispensed.

3. Necessity of Patent Protection For Ai Generated Work

The essential goal of IP framework is to empower advancement through new advances and inventive components which incorporates human creation just as advancement made and created by AI. Be that as it may, there emerges an uncertainty with regards to the responsibility for creation done by AI for example possession concerning both, the information just as the innovation that are the columns to any such creation. Notwithstanding, creation and advancement of new innovation is frequently trailed by insurance of Intellectual Property Rights (IPRs) and in the current situation, patent gives off an impression of being the best type of IP Protection.

If we want to understand the patentability of innovations created by AI, it is certifiable that the AI development is not a sig le development but there are a bunch of advancements. AI related innovations in the form of programming and the algorithm notwithstanding the equivalent fits the bill for patent security in India.

In India, the Patent Act, 1970 ('the Act') manages patentability of developments under the Act. Section 3(k) specifies the non-patentability of numerical and business techniques, PC programs essentially[14]. At the end of the day, there exists an unqualified boycott in India on the patentability of calculations and PC programs except if they mirror the modern appropriateness of the equivalent alongside oddity and non-conspicuousness of the advancement.

Another place of concern is that of non-conspicuousness of the development to the individual talented in the workmanship for example

[14] Section 3(k) of Indian patent Act, 1970, *available at*: https://indiankanoon.org/doc/141370947

in the field of AI, there is a probability that the advancement so made could seem, by all accounts, to be clear to the individual talented in the craftsmanship which may outcast the patent security of such a development. Further, according to the rule that everyone must follow, patent assurance is reached out to the first and genuine innovator and the designer should be a natural human.

In this way, it is just about as totally obvious light that AI can't be allowed theoretical rights and there comes the difficult inquiry of deciding the privileges of proprietorship with respect to the originator of such a thought for example the insight behind such an origin. For one to guarantee his development, one should add to its origination to be an innovator.

For making and building up a system of AI, human astuteness and inclusion is imperative. Computer based intelligence may act, carry on and learn like an individual be that as it may, as a trend-setter, common individual is needed for posting the name of an innovator.

Hence, there emerges an issue with respect to allow of patent for example regardless of whether patent can be allowed an "electronic individual" in this manner recognizing it from a "characteristic individual"? At the end of the day, an electronic individual might be considered as a designer and the lawful element or an organization attributing responsibility for same, might be given to the organization. While the patent laws across the globe expands and presents assurance of patent just to a characteristic individual and not to an electronic or a legitimate individual.

4. Role of IPR in the field of robotics.

Nowadays the field of robotics have enhanced vastly. Robotics was invented with the purpose of only to increase the industrial process i.e. to reduce the manpower and to increase the efficiency and production. The innovation in the field of robotics are happening at a very high speed. It seems that the science fiction is coming to reality. It is exceptionally synergistic, research-concentrated, and is getting progressively intricate. The human activities especially in the industrial sector are requiring the use of robotics over the past few years. It is no big surprise that the interest for developments and innovations in the field of mechanical technology is high.[15]

With the variety of likely utilizations of this innovation, it implies that customers can hope to acquire a wide scope of mechanical technology related items, from medical care to military. It can assume an urgent part in completing errands that might be considered excessively risky to people or helping people with restricted portability.

The focal point of mechanical technology development is moving from modern computerization to further developed advanced mechanics including different innovative fields, entertainers and monetary areas. Accordingly, related IP and different methodologies to suitable profits from advancement venture are early stage; our comprehension of them is deficient. Additionally, perceiving the expansive extent of the advanced mechanics industry is significant for this investigation on the grounds that the enormous assortment of mechanical technology items. furthermore, their applications implies that there is no uniform IP technique for mechanical technology organizations, nor are perceptions and patterns identified with one section of the advanced mechanics industry essentially pertinent to different portions of the advanced

[15] development in the Mechanical technology, *available at*: https://www.mondaq.com/unitedstates/trade-secrets/951200/today39s-robotics-innovation-landscape-and-the-role-of-ip-in-the-field-of-robotics

mechanics industry. Some conditional discoveries on allocation systems do, be that as it may, arise based on the existing writing, information and experiences from industry specialists and robotics technology analysts.

The patent under the IP laws have become important in the robotics industry because the amount of capital that is required in the research and development procedure is more only after this the manufacturing of the robots begins. In reality, the huge pre-statistical surveying and improvement costs combined with moderate administrative leeway can make a setting wherein exploring mechanical technology organizations feel needed to turn to patent assurance to recover their venture. Missing this security, new entrants would have the option to enter the market, after the "trail has effectively been blasted", at a lower cost for innovative work and need to conquer less administrative hurdles.[16]

For developments discoverable through figuring out or other legitimate methods, patent security is ordinarily preferred over proprietary advantages. It is perceived that numerous mechanical technology organizations whose upper hand is seen to be refined programming planned to empower automated equipment gadgets may utilize programming that is so confounded it can't be effectively figured out, which is something generally accepted to be conceivable with programming based electro-mechanical gadgets. In spite of the fact that discouraging and barring contenders is every now and again an essential thought of mechanical technology new companies, another basic motivator for looking for patent insurance concerns the seen benefits to

[16] Cooper (2013), referring to Casey Nobile and C. Andrew Keisner, The IP Battle Continues for Robotics Companies: "*A Patent Attorney's Reprise of the VGo/InTouch Health Verdict and its Implications*", Robotics Business Review, *January 7, 2013.*

new businesses when looking for investments.[17] Thus, key advanced mechanics developments were oftentimes protected by their unique – regularly scholarly – innovator, who frequently likewise began a relating organization or effectively moved the IP to existing assembling firms.

Thus, key advanced mechanics developments were oftentimes protected by their unique – regularly -scholarly – innovator, who frequently likewise began a relating organization or effectively moved the IP to existing assembling firms.

In entirety, Robotics patent expanded unequivocally during the 1980s, as wide based computerization of processing plants thrived and advanced mechanics research was increase. Advanced mechanics related first filings generally quadrupled during this decade. All the more critically, these filings dominated patent filings from other innovative fields. Advanced mechanics portion of complete licenses expanded from .13% in 1980 to 0.53% in 1993. At that point, after a moderately level protecting movement during the 1990s and first 50% of 2000s, the move to further developed mechanical technology has given another lift to mechanical technology protecting which proceeds right up 'til today. In a time of expanding generally speaking protecting movement, advanced mechanics supreme patent filings generally multiplied and the offer expanded from .4% in 2004 to .6% in 2011.[18]

[17] See Keisner (2012); See additionally Eilene Zimmerman, "*Why More Start-Ups Are Sharing Ideas without Legal Protection"*, The New York Times, July 2, 2014 referencing that beginning phase financial backers are for the most part hesitant to consent to non-exposure arrangements and talking about the advantages of recording temporary patent applications around there. This inspiration for new companies to document patent applications was certified by a few heads of IP for advanced mechanics new businesses during casual meetings directed by this current report's creator regarding this report.

[18] The relevant data related to patent filing, *available at*: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_econstat_wp_30.pdf.

Since 2009, there has been an expansion in the quantity of allowed robot licenses distributed around the world. This is demonstrative of the way that there is a huge expansion in innovative work in the field of advanced mechanics and the significance of security of advanced mechanics Innovations. What's more, with this enormous scope increment comes the generous need to ensure advancements and innovation.

Having a powerful protected innovation portfolio gives a few benefits to huge or little organizations. With protected innovation rights set up, the advanced mechanics organization has a superior benefit of situating themselves with possible angel investors. It is additionally ready to secure serious dangers, fabricate a solid brand name, and shield the organization from protected innovation robbery. As innovation keeps on developing, it is normal that the requirement for protected innovation insurance will increase.

5. IP Challenges In AI & Appropriate Policies To Address Them.

There are various advantages of Artificial Intelligence as it is bringing change in the life of numerous people. The AI is also bringing change in the field of IPR. Apart from various benefits the approach of AI has additionally raised various issues and challenges in the field of IPR which need to be addressed sooner or later. Artificial intelligence has been combined from engineering and science with the programming fed in them so that they can act or perform like humans. After years of research and development and complexities faced by the humans the AI has evolved drastically and now can perform functions beyond the capacity of a man, now its main emphasis has shifted to increase the functional effectiveness.

There are numerous challenges that AI is facing some of them are as follows[19]:

- **Liability challenge**: On the off chance that Artificial Intelligence can make a topic, it merits thinking about that in certain circumstances it can likewise be expected to take responsibility. Man-made consciousness could be exposed to claims of infringement of copyright, proprietary innovations (trade secrets), or even information protection, these can be infringed if it investigates the speculation plans of a business or customizes large information to a commercial advertisement, which is done by auto-duplicating details from various sources. Apart from this AI is something which is fed by humans i.e. it works what is being programmed in it. So AI can produce 3-d pictures, craftwork as well as artworks and if the work is being used without authorization than it can be charged with copyright or trademark infringement.

Presently, an question emerges, regardless of whether a computerized reasoning machine who learns different things all alone and builds up a speedy and exact interaction of producing data could be blamed for patent encroachment for utilizing the ensured innovation without realizing that it has effectively been licensed. However, again the question rehashes that, who is capable or obligated for its encroachment of copyright?

[19] Challenges that AI is facing, *available at*:
https://www.wipo.int/wipo_magazine/en/2019/01/article_0001.html

There is a solitude that Artificial Intelligence could possibly make such innovations that are totally against the improvement of living souls. In such situations where computerized reasoning clients ought to have the option to predict the outcomes and results, or are liable for overseeing and caring man-made brainpower, at that point they could be expected to take responsibility. In any case, if Artificial Intelligence works self-governing and can work with no immediate programming, create anything through self-learning and go past consistency, at that point the obligation or duty can fall upon the man-made AI itself.

- **Legislative Issues**: It has been seen that the laws identified with Intellectual Property have changed and adjusted every now and then in light of their dynamic presence and persistent new improvements by people. There should be a prerequisite for changes in existing laws identified with Intellectual Property to direct the development exclusively made by an AI machine and to choose which work ought to stay in the public area. Additionally, the officials should institute those laws which help in to choose which gatherings ought to be qualified for such creation and recognized as the proprietors of Intellectual Property coming about because of the making of AI[20]. To beat this issue with respect to the guideline of AI, all nations need to perceive similar cutoff points and essentials for delivering AI and make enactment covering the cures and every country's administrative system. This is the best way to determine the debates identified with the Intellectual Property of AI machine.

[20] legislative issue, *available at*: https://corporate.findlaw.com/intellectual-property/current-copyright-issues.html

- **Challenges relating to policies**: The logical distributions and patent information are proof of the fast speed of AI development. This pattern has been joined with the utilization of numerous AI advancements and their future effect on the regular day to day existences of individuals, implies AI innovations represent an arrangement challenge for government, policymakers, and controllers. These difficulties incorporate the insurance of individual information of each resident, the improvement of guidelines and standards identified with information sharing, how development can be supported, the guideline of new man-made reasoning advances, and furthermore to save people from the danger started from exceptionally progressed man-made consciousness machines.

- **IP infringement**: The inquiry is if AI can claim an Intellectual Property, regardless of whether it can encroach the privilege of the outsider or not? Or on the other hand we can say that if AI machine is recognized as having the option to deliver the topic which was at that point there, presently who might be expected to take responsibility if that topic encroaches the Intellectual Property of the outsiders? The issue is particularly pertinent with regards to an encroachment of copyright which includes real duplicating from the protected source, that is, the creator of a work which encroached the copyright of an outsider probably approached that protected work. On the off chance that we talk about the circumstance identified with a AI machine, in which we as a whole expect that a AI machine should approach everything accessible on the web, the test of demonstrating that the infringer approached the secured work will be a lot simpler to address.

Ownership challenge: If one wants to protect his/her piece of work under copyright law the work should have some skill and judgement and must be unique. In like manner, the author has the privilege as he is the primary proprietor of a plan he made, and the innovator is the principal owner of a patent. In these conditions, the proprietor is straightforwardly identified with the topic being created. An innovation should incorporate materialness, curiosity, and creative strides for a patent to be allowed. The laws identified with licensed innovation face a significant test when the councils attempt to respond to an inquiry, regardless of whether computerized reasoning has made sufficient strides or uses these variables to produce a piece started after a long exploration work or not.

Appropriate policies or solutions to overcome these challenges

Regardless of the discussion and difficulties, there is a need to create a functional and proper approach to manage the current circumstance. The outcomes produced by AI are either its very own result insight or a calculation. On the off chance that the AI machine's capacity is simply mechanical, instead of imaginative, AI may be considered to need inventiveness. The current Intellectual Property law of any nation would not distinguish Artificial Intelligence as a proprietor of Intellectual Property. Hence, AI innovation would not have the option to get proprietorship for any creation until and except if it's ready to accomplish the lawful character status like people. Most nations' current Intellectual Property laws necessitate that any rights holder should have a legitimate character that is deficient on account of Artificial Intelligence. As we know Saudi Arabia has become the 1st country to grant citizenship to the robot named Sophia.[21]

[21] Information related to Sophia, *available at*:

Sooner rather than later, Artificial Intelligence can before long rise above human knowledge and lead the general public to revelations, and to accomplish these objectives, the Intellectual Property law of the nation should defend the interest of Artificial Intelligence advancements. In the event that AI can demonstrate a free development, it very well may be viewed as a likely creator and might guarantee a copyright for such advancement as other human creators. AI machines equipped for improving and broadening their abilities through learning and preparing may be equipped for patent possession for the advancement they produce.

There is a worry that man-made reasoning could possibly make such developments that are totally against the advancement of living souls. In such situations where AI clients ought to have the option to anticipate the outcomes and results, or are liable for overseeing and really focusing on man-made consciousness, at that point they could be expected to take responsibility. Regardless, if AI works self-sufficiently and can work with no immediate programming, create anything through self-learning and go past consistency, at that point the obligation or duty can fall upon the AI itself. The enactment ought to be drafted so that will guarantee that individuals reserve the privilege to supersede any man-made brainpower choice and furthermore award capacity to people to control the working of man-made reasoning.

In the event that we investigate the fundamental destinations of the Intellectual Property law, the vital approach of these laws exists to give select rights to financial backers or makers to appreciate the advantages given by their separate works. In the event that similar rights were given

https://www.dw.com/en/saudi-arabia-grants-citizenship-to-robot-sophia/a-41150856 to AI advancements, they would likewise have the option to perceive the accomplishment they brought about from such creation or by concocting any topic and would appreciate the advantages. In any case, offering worth to advancement that gives advantage to the overall population is a crucial objective of the Intellectual Property law, and barring such development from giving rights would be conflicting with the Intellectual Property law. It is conflicting to the arrangement identified with Intellectual Property law if the administrator thought about innovativeness and advancement over the improvement of individuals. There should be a sensible harmony between these two conditions.[22]

Conclusion

The researcher hereby concludes that with the growth of Artificial intelligence there hasn't been the growth in the IP laws. With the pace the technology is going to another level, so there is a need that the law relating to IP should get amended and the new provision must get added so that the work of the owner gets protected. In Saudi Arabia, a humanoid named Sophia, a computerized reasoning robot has acquired rights that are accessible to the residents of Saudi Arabia and has been allowed citizenship of Saudi Arabia in 2017. Sooner rather than later, AI will get affirmation for their commitments to society. Sooner rather than later, AI dramatically affects being human and even can perform what ordinary people can't. There should be an appropriate harmony between the use and commercialization of the advancement made by man-made consciousness so it can give sensible advantages to the man-made reasoning machine and should give advantages to general society everywhere or we can say that it should be of public interest.

[22] https://www.lexology.com/library/detail.aspx?g=d5acda9a-7e17-4a0e-b9a1-34bd4a8b4248

The coming of 21st century, has stamped and set up the broad utilization of electronic devices in each and every article that has created across the globe. Directly from moving toys to anthropological robots and from PC controlled vehicles to atomic reactors, next is making gadgets and innovation insightful and smart. This uncommon mix of concurrent presence of human and innovation as AI, will go about as a defining moment in the monetary improvement of the world on the large. Along these lines, it is the need of great importance wherein recent concerns in regards to AI should be tended to in the best way. The current conflicting site of AI instrument comparable to IP laws worries about the need to perceive the manifestations and creations made out of AI framework. Besides, expecting the elements of AI frameworks with no human intercession requires a solid insurance of such developments.

Accordingly, a uniform treatment of the AI system will unquestionably go about as a positive advance and go about as a rousing variable being developed of new creations. All the part countries of multilateral arrangements should perceive the confusing circumstance of AI instrument, by carrying a correction to TRIPS.

In any case, the genuine issue lies at a point where the execution of AI instrument versus IP framework is to be offered impact to however on the off chance that the equivalent is executed with an all-around arranged methodology, it will unquestionably give a jump in the field of developments in AI and IP laws soon.

It is obvious there is a massive gap between ground reality and existing regulations with too many challenges brought on by AI. Today, IBM has the largest portfolio of AI related patent applications with 8,290 patent applications in the world, followed by Microsoft with 5,930 patent applications[23]. With this increasing popularity of AI related inventions and the sheer volume of AI related patent applications being filed, it will be up to the patent offices and regulators to revise existing patent and IP laws and create new molds to fit the emerging technology.

[23] Patent application, *available at*: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf

PREVENTING CRIMINAL MISUSE OF INNOVATION USING AI

Dr. Anuradha Girme[1] and Utpal Gharde[2]

Abstract

Technological advancement has changed industries and made daily life easier. But it has also created new ways for criminals to misuse these technologies. Cyber fraud, deepfake manipulation, AI-based phishing, and large-scale data breaches are becoming more common. Criminals are using modern tools to commit such acts. As these threats become more complex, AI is now a key tool to stop the misuse of technology.

This paper studies how AI is being used to prevent technology-based crimes. It looks at four main areas: cybersecurity, financial fraud detection, social media monitoring, and law enforcement. AI systems improve cybersecurity by finding unusual activities, detecting threats, and responding quickly. Machine learning models can study large amounts of data in real time. They help detect cyberattacks, malware, and unauthorised access.

In finance, AI is used to spot fraud, money laundering, and identity theft by studying behaviour patterns. Deepfake detection tools, using digital forensic methods, help identify fake content. Social media platforms use AI to control cyberbullying, false news, and hate speech.

Even though AI is helpful, it also creates legal and ethical problems. Major concerns are about data privacy, bias in algorithms, and lack of openness. These problems need strong rules and systems that respect rights and follow the law.

[1] Assistant Professor, New Law College, Bharati Vidyapeeth

[2] LLM (2nd Yr), New Law College,

This study says that AI can protect digital spaces, but only if it is used in a fair and responsible way, keeping people's rights in mind.

Keywords: *Ethical AI governance, Technology-facilitated crime, AI and digital forensics, Cybercrime prevention, Privacy and surveillance*

1. **Introduction**

Technology has grown rapidly in the twenty-first century. This growth has brought major digital changes across the world. It has reshaped economies, changed how governments work, and affected the way people interact every day.

At the centre of this change is AI. AI now plays an important role in key areas like healthcare, finance, transport, and public services. It can automate decisions, handle large amounts of data, and copy human behaviour. These features have made AI a key part of today's innovation.

However, this progress also brings new risks. AI can be misused for criminal purposes. As AI becomes more powerful and easier to access, criminals are using it in harmful ways. Hackers and cybercrime groups now use AI to carry out phishing, identity theft, spread false information, and commit financial fraud. Generative Adversarial Networks (GANs) are used to make realistic deepfakes. [3] AI bots help with social engineering and carry out cyberattacks. [4] These activities increase the reach and secrecy of crimes and reduce public trust in digital systems.

[3] Socradar® Cyber Intelligence Inc., *The Top 10 AI Tools for Deepfake Detection in 2025* (Mar. 1, 2025), https://socradar.medium.com/the-top-10-ai-tools-for-deepfake-detection-in-2025-8397a2ca8c22

[4] Max Smeets, *How AI Will Change the Future of Cyber Operations*, 72(1) *Survival* 27 (2020).

To deal with these risks, governments, companies, and international bodies are starting to use AI to prevent crime. It is being applied in areas such as predictive policing, fraud detection, fake media identification, and content moderation on social media.[5] But this also brings legal and ethical concerns. Issues like privacy, bias in AI, lack of clarity in decisions, and possible violation of legal rights are now widely discussed in research and policymaking.

This paper studies the two sides of AI. It looks at how AI is used to commit crimes and also how it helps to stop them. The paper uses doctrinal legal research to study this. It focuses on four main areas: cybersecurity, financial fraud detection, deepfake control, and social media monitoring.

The research aims to answer three main questions:
How is AI being used to prevent digital crime in different sectors?
What legal and ethical problems come up when AI is used in crime prevention?
What policies are needed to ensure AI is used in a fair and rights-based manner?
The study looks at legal and policy developments in Western countries and global organisations. It closely studies laws, policies, and expert writings. The aim is to build a balanced legal approach that allows the safe use of AI while protecting key legal rights and principles.

2. **Literature Review**

The relationship between Artificial Intelligence (AI) and crime prevention has received growing attention in recent academic and policy literature. Scholars

[5] Rashida Richardson, Jason M. Schultz & Kate Crawford, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 *N.Y.U. L. Rev. Online* 15 (2019)

agree that AI plays a dual role. It can support law enforcement efforts, but it can also be misused by criminals. This dual use is widely discussed in legal, technological, and governance studies.

Several recent publications have highlighted the increasing use of AI in cybercrime. Europol's Serious and Organised Crime Threat Assessment 2023 reports that AI enables more targeted, scalable, and low-risk offences such as phishing, ransomware, and deepfake creation. Scholars such as Smeets (2020) and Ajder (2019) have noted that criminal actors now use tools like Generative Adversarial Networks (GANs) to generate realistic synthetic media. However, many countries lack the legal capacity to respond. Legal frameworks in developing and even developed jurisdictions are often outdated or incomplete.

In the cybersecurity domain, literature shows both promise and concern. Academic and technical reports support the view that machine learning models—supervised and unsupervised—improve detection of cyber threats. These are commonly used in systems like Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS). However, legal scholars such as Hildebrandt (2018) raise concerns about privacy and data protection under laws such as the General Data Protection Regulation (GDPR).

In the financial sector, studies highlight AI's role in fraud detection, anti-money laundering (AML), and risk scoring. Case studies like JPMorgan Chase's COiN platform are often cited as successful examples. Yet, scholars like Angwin et al. (2016) and Selbst (2019) warn of bias and lack of transparency in financial algorithms. False positives and discriminatory outputs continue to affect marginalised communities.

Research on deepfakes and digital forensics is also expanding. According to Farid (2020) and Rössler et al. (2019), detection tools like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are being developed to spot manipulated content. Despite technical progress, the literature notes that detection must improve rapidly to match the pace of innovation in generative tools.

Social media governance is another area of focus. Scholars such as Douek (2021) and Lynch (2019) have studied the role of AI in content moderation. Natural Language Processing (NLP) tools are used to detect hate speech, misinformation, and cyberbullying. However, the literature warns against over-censorship, cultural misinterpretation, and lack of user redress mechanisms. These concerns are particularly important in the context of constitutionally protected speech and international human rights standards.

Common ethical concerns appear across all sectors. These include data privacy, user consent, algorithmic bias, and opacity. Authors like Wachter and Mittelstadt (2021) argue for stronger safeguards and human rights-based governance. However, the literature agrees that existing legal systems have not fully translated ethical principles into enforceable laws.

3. **Methodology**

This research follows a doctrinal legal approach. It involves a detailed study of legal texts such as statutes, court decisions, regulations, and academic writings. The aim is to examine how AI is used in crime prevention. The study focuses on understanding current legal frameworks, finding gaps, and giving legal and policy suggestions. These suggestions are based on constitutional values and ethical principles.

Source Selection

This study uses both primary and secondary legal sources.

Primary sources include:

The General Data Protection Regulation (GDPR), the UK Data Protection Act, and U.S. state-level AI laws.

Court judgments on topics like surveillance, algorithmic bias, and digital forensics.

Official reports from law enforcement and regulatory bodies. These include Europol's threat assessments and policy documents from the European Commission, Interpol, and the UN Office on Drugs and Crime.

Secondary sources include:

Peer-reviewed journal articles and legal commentaries on technology law, AI ethics, and criminal justice.

Reports from tech companies such as Microsoft and Google on deepfake detection and AI regulation.

White papers by banks like JPMorgan Chase on how AI is used in anti-fraud systems.

Publications from human rights organisations and research centres on AI's legal and social impact.

Analytical Framework

The research combines legal reasoning, content analysis, and case study review.

Legal reasoning is used to study laws and court decisions about data protection, surveillance, and discrimination.

Content analysis is used to review how academic and policy discussions explain the risks and benefits of AI.

Case studies show how AI works in the real world. Examples include Europol's use of AI to study organised crime, JPMorgan's COiN system to detect fraud, and AI-based content moderation on platforms like YouTube and Facebook. These cases help identify legal and regulatory challenges in actual AI use.

Scope and Limitations

The research looks at four areas where AI is widely used and where risks are high: Cybersecurity, financial fraud detection, deepfake identification, and social media monitoring. These are areas where AI is both misused by criminals and used to prevent crime.

The study focuses on Western countries—mainly the European Union, United Kingdom, and United States—as these places have strong rules for AI and data protection.

While the paper explains how AI is used in practice, it does not go into the technical side of AI models or algorithm design. The focus stays on legal aspects. Also, since legal systems differ across countries, the findings may not apply everywhere—especially in countries without strong AI or privacy laws.

By using legal methods and studying current rules, this approach helps maintain both legal clarity and ethical relevance. It allows a clear understanding of how the law can deal with fast-growing technology while still protecting people's rights.

4. Findings

This section presents the main findings of the study. It is organised around four important areas where AI is currently used to fight technology-based crime. The findings are based on case studies, legal analysis, and institutional reports. Each area shows the strengths of AI, as well as its limitations.

4.1 Cybersecurity

AI has made cybersecurity stronger by helping in both early warning and real-time threat detection. Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) now use both supervised and unsupervised machine learning models. These systems track network traffic, spot unusual behaviour, and send alerts without delay.

For example, if a company server sends encrypted data outside working hours, the AI can spot it and lock down the system. This prevents damage before the threat grows.

AI also improves how quickly organisations respond to attacks. Automated platforms collect and study data from many sources—like firewalls, emails, and log systems. Once a threat is found, the AI can isolate the affected device or update firewall settings.

MIT CSAIL's AI combines human input with AI prediction and has shown over 85% accuracy in detecting cyber threats.[6]

4.2 Financial Fraud Detection

Banks and other financial institutions are now using AI to detect fraud, identity theft, and money laundering. AI tools study user behaviour and look

[6] A. Neupane, N. Saxena & S. Chattopadhyay, *Explainable Intrusion Detection Systems (X-IDS): A Survey*, 113 Computers & Sec. 102577 (2022), https://doi.org/10.1016/j.cose.2021.102577

for abnormal activities.[7] For example, if a person makes quick withdrawals from different cities or sends money abroad suddenly, the system raises an alert.

AI is also used in biometric checks.[8] Deep learning helps power facial recognition, fingerprint matching, and voice authentication. These systems use “liveness detection” to confirm that a real person—not a fake input or deepfake—is present.

JPMorgan Chase’s COiN platform uses Natural Language Processing (NLP) to study contracts and transactions. This reduces errors and saves time. In anti-money laundering (AML) systems, AI maps risky account networks and tracks illegal money flows.

4.3 Deepfake Detection

AI plays an important role in spotting and stopping deepfakes. Detection tools use Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to check images and videos. They find signs like odd pixels, strange facial expressions, or mismatched lip movements.

AI can also detect fake audio. It checks for unusual speech patterns or sound distortions.[9] Microsoft’s Video Authenticator is one such tool. It gives each video frame a score to show if it might be fake.

[7] Reuters, *JPMorgan’s AI System Helps Boost Sales, Expand Client Base, and Save $1.5 Billion in Fraud Prevention* (May 5, 2025), https://www.reuters.com/business/finance/jpmorgans-ai-system-saves-15-billion-fraud-losses-2025-05-05/.

[8] Jennifer Lynch, *Face Off: Law Enforcement Use of Face Recognition Technology*, Electronic Frontier Foundation (EFF), https://www.eff.org/wp/law-enforcement-use-face-recognition.

[9] Hany Farid, *Digital Forensics in a Post-Deepfake World*, 12(6) *IEEE Signal Processing Magazine* 82 (2020).

Big datasets like Google's Deepfake Detection Challenge and FaceForensics++ are used to train these systems. [10] These tools could be used in courtrooms and by journalists to confirm whether digital content is real.

4.4 **Social Media Monitoring**

AI-based Natural Language Processing (NLP) tools are widely used to check online content. [11] These tools detect hate speech, cyberbullying, false news, and signs of radical thinking. They work across many languages and analyse tone and emotion. They also flag violent or harmful content.

Platforms like Facebook and YouTube say that their AI removes substantial harmful posts before users report them.

AI is also used to study behaviour patterns related to self-harm or recruitment into extremist groups. On the dark web, AI scans and classifies illegal material. Police use it to track drug trade, human trafficking, and illegal markets.

Graph-based machine learning models help find links between anonymous users. These models reveal how online crime groups are connected.

These findings show that AI is playing a bigger role in protecting digital systems and improving public safety. However, as the next section will explain, there are also new risks and legal gaps that need urgent and careful action.

[10] Andreas Rössler et al., *FaceForensics++: Learning to Detect Manipulated Facial Images*, arXiv preprint arXiv:1901.08971 (2019), https://arxiv.org/abs/1901.08971.

[11] Europol Innovation Lab, *Guidelines on Ethics for AI in Law Enforcement Video Analytics* (2023), https://www.europol.europa.eu/publication/ethical-ai-video-guidelines.

5. Analysis

The findings in the previous section confirm that AI is now essential in the fight against technology-facilitated crimes. However, its deployment in areas such as cybersecurity, financial services, media verification, and content moderation presents several legal and ethical challenges. This section critically examines those concerns in the context of fundamental legal principles and human rights protections.

5.1 Legal Effectiveness vs. Data Privacy

AI's strength lies in its ability to process vast quantities of data, enabling real-time threat detection and rapid situational response. Yet, this capability often depends on large-scale data collection and surveillance. Such practices may conflict with privacy rights protected under laws like the General Data Protection Regulation (GDPR).

For instance, Intrusion Detection Systems (IDS) that track user activity on networks may capture personal and behavioural data without explicit consent. [12] Although GDPR allows data processing for legitimate security interests, it also mandates data minimisation and transparency.

Many AI systems operate as "black boxes". [13] This means individuals cannot easily determine how their data is processed, how decisions are made, or whether their rights are respected. This lack of transparency undermines the right to informational self-determination and weakens legal accountability.[14]

[12] Peter Parycek et al., *Data Protection and Privacy in Intrusion Detection Systems*, 33(1) *Int'l Rev. L. Comput. & Tech.* 85 (2019).

[13] Sandra Wachter, Brent Mittelstadt & Chris Russell, *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, 41(4) *Comput. L. & Secur. Rev.* 105567 (2021).

[14] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* 19–21 (Harvard Univ. Press 2015).

5.2 Bias and Discrimination in Automated Decision-Making

AI systems used in fraud detection and law enforcement risk reproducing structural biases embedded in historical data. If training datasets reflect social inequalities, the resulting models may produce skewed outcomes.

For example, credit-scoring algorithms have flagged marginalised groups as high-risk, while facial recognition tools often perform poorly on non-white faces—a concern documented in multiple studies.[15] These outcomes may breach principles of substantive equality and due process, especially in contexts that affect financial access or criminal investigations.[16]

Furthermore, the opacity of many AI systems makes it difficult for individuals to challenge automated decisions. Without clear explanations, affected persons may lack meaningful ways to seek redress or contest discriminatory outcomes.

5.3 Deepfake Detection and Evidentiary Integrity

Deepfake technologies pose serious threats to the credibility of digital evidence. This is particularly concerning in criminal trials and public discourse. AI-based detection tools help, but their reliability depends on constant updates and probabilistic scoring.

Legal systems that rely on proof beyond reasonable doubt may hesitate to accept AI-authenticated media unless forensic standards are clearly met. In addition, generative tools are now widely accessible. This raises concerns for

[15] Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 *Proc. of Machine Learning Research* 77 (2018), https://proceedings.mlr.press/v81/buolamwini18a.html.

[16] Europol, *Understanding and Mitigating Bias to Harness AI Responsibly* (June 26, 2023), https://www.europol.europa.eu/media-press/newsroom/news/understanding-and-mitigating-bias-to-harness-ai-responsibly.

freedom of expression, especially if detection tools wrongly censor satire or political speech.[17]

Such tensions call for robust legal frameworks. These should clarify what constitutes lawful creation, detection, and use of synthetic media, while balancing security with free expression.

5.4 Censorship and Surveillance Overreach

AI-driven content moderation has improved the detection and removal of harmful content. However, it also increases the risk of overreach. In opaque or state-controlled environments, such systems may be used to suppress dissent.[18]

There have been reports of disproportionate takedowns targeting specific political or ethnic groups. This raises concerns about AI being weaponised for digital repression. On the law enforcement side, AI tools used to monitor the dark web often operate without judicial oversight. Practices such as bulk data collection and warrantless surveillance challenge the principle of reasonable privacy.

The lack of transparent oversight mechanisms—whether judicial, parliamentary, or independent—further amplifies the risk of abuse.

5.5 Fragmented Regulation and the Need for Legal Coherence

AI governance remains fragmented across jurisdictions. While the European Union has introduced the AI Act and strengthened GDPR enforcement, many other countries lack comprehensive laws.[19]

[17] Evelyn Douek, *Governing Online Speech: From 'Posts-As-Trumps' to Proportionality and Context*, 121 *Colum. L. Rev.* 1 (2021).

[18] Shoshana Zuboff, *The Age of Surveillance Capitalism* 413–422 (PublicAffairs 2019).

[19] European Commission, *Proposal for a Regulation Laying Down Harmonised Rules on AI (AI Act)*, COM/2021/206 final (Apr. 21, 2021).

This patchwork approach weakens regulatory coherence. [20]It also complicates international efforts to address cross-border cybercrime. A more harmonised legal framework is needed—one grounded in human rights, democratic accountability, and the rule of law.[21]

Key concepts such as Explainable AI (XAI), human-in-the-loop decision-making, and algorithmic auditing should become standard in public-sector AI deployments. [22] These principles can help ensure that AI systems serve justice rather than erode it.

This analysis reveals a fundamental tension. While AI enhances the ability to detect and prevent crime, it also exposes the limitations of existing legal structures.

6. Policy Recommendations

To ensure that AI is used both effectively and ethically in preventing criminal misuse, a strong policy framework is necessary. The following suggestions are based on legal principles and the findings of this study. They aim to balance technology, human rights, and democratic checks.

6.1 Set up Independent Regulatory Authorities

Governments should create independent regulators to oversee the use of AI in law enforcement and public safety. These bodies must have legal powers and full independence. Their main roles should include:

[20] Inter-American Development Bank, *Cracking Crime with AI* (2025), https://www.iadb.org/en/news/cracking-crime-ai.

[21] Organization for Economic Co-operation and Development (OECD), *Recommendation of the Council on AI*, OECD/LEGAL/0449 (2019), https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[22] Finale Doshi-Velez & Been Kim, *Towards a Rigorous Science of Interpretable Machine Learning*, arXiv preprint arXiv:1702.08608 (2017), https://arxiv.org/abs/1702.08608.

Certifying AI systems used by public and private sectors. Certification must check for accuracy, fairness, openness, and respect for human rights.

Conducting regular audits of algorithms. High-risk tools like biometric surveillance, credit scoring, and predictive policing must face stricter checks.
Investigating wrong use of AI and punishing violations, including biased or illegal applications.

The European Union's AI Act can be used as a model. But each country should adapt it based on its own legal and constitutional values.

6.2 Create Legal Frameworks for Ethical AI Use

Ethical AI principles must become part of binding laws. Voluntary guidelines are not enough. AI ethics laws should include:
Human Rights by Design: AI tools must protect privacy, equality, and due process from the start.[23]
Risk-Based Rules: Classify AI tools by how risky they are. High-risk tools should follow stricter rules.
Public Involvement: Laws must be made with input from civil society, affected people, legal and ethical experts, and technology professionals.
These frameworks should be legally enforceable to ensure public trust and accountability.

6.3 Make Explainable AI and Human Oversight Mandatory

Explainable AI (XAI) should be compulsory in sensitive areas. Laws and policies must ensure:
AI models produce results that ordinary users can understand.

[23] U.N. Human Rights Council, *The Right to Privacy in the Digital Age*, U.N. Doc. A/HRC/27/37 (2014).

Human decision-makers must be involved. They should be able to check, question, or override AI decisions.

Developers must clearly document training data, how the model works, its limits, and test results.

Explainability is important for protecting user rights, ensuring fair trials, and allowing courts to review AI-based decisions.

6.4 Build Public–Private Partnerships (PPPs)

Private companies have deep technical knowledge. [24]Governments must build safe and clear partnerships with them. These PPPs should:

Share information about cyber threats, fraud, and deepfakes while protecting user privacy.

Test AI tools in controlled settings before launching them.

Help set ethical standards and train police and other officials to use AI responsibly.

Clear legal rules must control what data can be shared and how these partnerships are supervised.

6.5 Improve Public Awareness and Professional Training

A well-informed public and trained professionals are key to good AI governance. Policy frameworks should:

Include digital and AI education in schools and universities. Focus must be on privacy rights and how algorithms work.

Provide regular training for police, judges, and policymakers. They should learn how to read AI results, find bias, and protect legal rights.[25]

[24] Tim O'Reilly, *What Is Platform Strategy?*, Harvard Business Review (May 27, 2020), https://hbr.org/2020/05/what-is-platform-strategy.

Support courses that combine law, ethics, and technology. This will help create experts who understand both legal and technical issues.
People and professionals must be aware of how AI works to ensure responsible use.

These policy steps support the ethical use of AI in legal and enforcement systems. With proper laws, ethical design, open institutions, and public involvement, AI can be used to protect rights and strengthen justice.

7. Conclusion

AI has become a powerful tool. It improves human abilities but also changes how crimes are planned and committed. This study has looked at AI's two-sided role. It can be misused by criminals, but it is also useful for stopping crime. The paper studied AI use in areas like cybersecurity, financial fraud detection, deepfake control, and social media monitoring. The results show that AI helps detect, study, and respond to digital threats quickly and accurately.

However, these benefits come with risks. The same qualities that make AI strong—like autonomy, use of large data, and flexibility—can also harm people's rights. Problems such as mass surveillance, bias in decisions, lack of transparency, and weak legal checks show why strong rules are urgently needed. If not properly controlled, AI may increase inequality, reduce privacy, and harm justice systems.

This research answered three main questions:
How is AI used to prevent technology-based crimes?

[25] Karen Yeung, *Algorithmic Regulation: A Critical Interrogation*, 12(4) *Regulation & Governance* 505 (2018).

→ AI is used in many ways. These include detecting suspicious behaviour, verifying identity through biometrics, checking media for fakes, and removing harmful online content. These tools are already part of daily work in security and law enforcement.

What legal and ethical problems come from using AI for crime prevention?→ there are major concerns. These include risks to privacy, unfair decisions, unreliable digital evidence, and systems that people cannot understand or question—especially in areas like policing and finance.

What policies can ensure responsible use of AI?
→ Good AI use needs strong rules, independent oversight, ethical design, human involvement in decisions, and public education. These steps must go together to match technology with human rights and legal values.

Looking forward, AI laws and policies must focus on transparency, fairness, and inclusion. AI should not be judged only by how fast or smart it is. It must also support democracy and the rule of law. Future research should check how well AI laws work over time, how Explainable AI (XAI) can be used in practice, and how countries can work together to manage cross-border AI issues.

AI is not good or bad by itself. It reflects the goals and values of its makers and users. If used wisely, AI can be a strong partner in fighting crime. But if left unchecked, it can cause more harm than good. The real challenge is to build legal, ethical, and institutional systems that make sure AI protects the same rights it could otherwise put at risk.

CYBER WARFARE IN THE LENS OF INTERNATIONAL LAW: LEGAL COMPLEXITIES AND PATHWAYS TO REFOR

Dr. Shivangi Sinha & Ms. Niha Khan [1]

Abstract

Cyber warfare has emerged as a critical aspect of modern conflicts, raising significant legal, ethical, and security concerns. As cyberspace becomes a battlefield for state and non-state actors, understanding the applicability of international laws, particularly International Humanitarian Law (IHL), is crucial. This research paper provides a comprehensive analysis of cyber warfare, beginning with an exploration of cyberspace and its role in modern conflicts, tracing the historical evolution of cyber warfare, and identifying key actors involved in cyber operations. It further examines the existing international legal frameworks governing cyber warfare and assesses their effectiveness in addressing emerging threats.

A detailed analysis of cyber warfare incidents and their repercussions highlights the challenges posed by such operations, including their impact on national security, civilian infrastructure, and global stability. The study also evaluates the applicability of IHL to cyber warfare and identifies gaps within the current legal framework. In light of these challenges, the paper proposes necessary legal reforms to ensure that international law remains relevant and effective in regulating cyber operations.

The information presented in this research paper has been gathered from a wide range of credible sources, including academic books, peer-reviewed journal articles, legal reports, and authoritative websites. By offering a

[1] Assistant Professors of Bharati Vidyapeeth (Deemed to be University), New Law College, Pune

thorough examination of the intersection between cyber warfare and international law, this study aims to contribute to the ongoing discourse on strengthening legal mechanisms to address cyber threats effectively.

1. **Introduction**

The internet has made the world more globalised and interconnected, creating an environment in which organisations rely on data flows to execute their daily operations, influencing everything from operability to business models. According to a 2017 article in The Economist, data is becoming as important to society as oil due to its interconnected nature and inability to be disconnected.[2]

Cyberspace refers to the "virtual" environment generated by links between computers, Internet-enabled devices, servers, routers, and other Internet infrastructure components. In contrast to the Internet itself, cyberspace is a place created by these links. Some believe it exists independently of any nation-state. The term "cyberspace" was first used by American-Canadian author William Gibson in 1982 in a tale published in Omni magazine, followed by his book Neuromancer. [3] As stated in the doctrine (Melzer, 2017), cyberspace is the "fifth sphere or fifth domain of warfare" following land, sea, air, and outer space. This argument is unarguable since, due to the advancement of current technologies, cyberspace is a possible theatre of military operations.[4]

[2] Mikael Weissmann et al. eds., Hybrid Warfare: Security and Asymmetric Conflict in International Relations (Routledge 2021).

[3] Cyberspace, Encyclopaedia Britannica, https://www.britannica.com/topic/cyberspace (last visited Apr. 7, 2025).

[4] Garkusha-Bozhko, International Humanitarian Law in Cyberspace: Ratione Materiae, Ratione Temporis and the Problem of Qualification of Cyberattacks, 3 J. Pol. & L. 45 (2021).

There is a difference between the terms cyber-attack, cyber-crime, and cyber warfare.

Cyber-attack, according to Waxman, is any attempt to alter, disrupt, or destroy computer systems, networks, or their associated information or programs. The German Cyber Security Strategy defines a cyber-attack as "an IT attack in cyberspace directed against one or several other IT systems aimed at damaging IT security—confidentiality, integrity, and availability—which may be all or individually compromised."

Cybercrime is simply described as "any crime that is facilitated or committed using a computer network or hardware device." Cybercrime encompasses a wide range of criminal actions in cyberspace, such as cyber-squatting, online privacy violations, and the storing and broadcasting of child pornography, among others.[5]

Cyber warfare is a digital battle that uses information and communication technology to target the security of an attacked state and do severe damage. Cyber warfare differs from conventional armed battles. "The most important characteristic of cyber warfare is that it takes place partially or entirely in cyberspace or through it (by acting from cyberspace on the physical world and vice versa)." Adkins defines digital warfare as "the use of computer techniques of intrusion and other capabilities against the opponent's infrastructure based on information and communication technologies, with the intention of compromising national security or preparing for future operations against national security." [6]According to Martin C. Libicki, there are two sorts of cyberattacks: strategic and operational. The former entails a

[5] Hemen Philip Faga, The Implications of Transnational Cyber Threats in International Humanitarian Law: Analysing the Distinction Between Cybercrime, Cyber-Attack, and Cyber Warfare in the 21st Century, 16 Balt. J.L. & Pol. 144 (2023).

[6] Sanela Veljković, Possibility of Applying the Rules of International Humanitarian Law to Cyber Warfare, 41 Pravo - teorija i praksa 17 (2024), https://doi.org/10.5937/ptp2403017V.

campaign of cyberattacks against another entity, while the latter involves using cyberattacks against the other side's military during a physical battle. An actor may use malicious code to attack or harm an adversary's computer or network, attacking online control systems that manage both logical and physical networks. [7]

1.1 The Historical Evolution of Cyber Warfare

Although cyberspace and cyber-attacks are a new phenomenon, attacks on data and databases are not. In fact, technically the first cyber-attack took place in France even before the internet came into existence in the year 1834 when attackers accessed the French telegraph system and stole financial market information. In the 20th century after the advent of the internet and the emergence of cyberspace, cybercrimes became its own ecosystem and further evolved into large scale cyber warfare. One of the first large-scale cyber-attacks took place in 1988 when a Cornell Student created a virus called 'Morris Worm' which infected approximately 10% of the computers connected to the internet at that time.[8]

Cyberspace became a prominent war field for the first time during the cold war. Russia's cyber-attack on the US during the late 1990s was one of the first state sponsored cyber-attacks. Evidentiary details regarding the attack are less in the public domain because the details are still kept classified due to the nature of the stolen information and due to the ground-breaking significance of the attack. The attack referred to as the 'Moonlight Maze' attack started in 1996 went for 2 years unnoticed and is one of the longest cyber-attacks in history. The attacks which hit various US governmental departments

[7] Anne-Marie Eklund Löwinder & Anna Djup, Cyberwarfare and the Internet: The Implications of a More Digitalized World, in Hybrid Warfare: Security and Asymmetric Conflict in International Relations 145 (Mikael Weissmann et al. eds., I.B. Tauris 2021).

[8] Christopher Kelty, The Morris Worm, 1 Limn (2011), https://escholarship.org/uc/item/8t12q5bj.

including the Pentagon and NASA stole information such as troop configuration, maps of military installations and military hardware designs. The task force that investigated the case claimed that if all the stolen material was printed and stacked it would have been 3 times the height of the Washington monument[9].

The 2000s gave rise to 'hacktivism' where politically motivated individuals or groups would hack into sensitive or classified databases. Today cyberspace has evolved into sophisticated and crucial weapons used by nation-states, terrorist organisations and even disgruntled individuals.

1.2 Key Actors in Cyber Warfare

In today's day and age there are multiple threat actors in cyberspace. They can be categorised based on their method of operation, goals and motivation.

- Nation-State Actors- Malicious actions are carried out by nation-state threat actors on behalf of a particular government or nation-state. Professional hackers are frequently employed to carry out targeted assaults against other nations or organizations. The majority of nation-state actors have economic or political motivations. They will be well funded and have better technological infrastructure compared to individual actors and they engage in serious violations such as espionage, election interference or gaining access to critical government information. Nation-state actors breach government security systems using a variety of advanced tactics. Many state-sponsored cyber threat actors, for instance, develop into advanced persistent

W. Gragido & J. Pirc, The Rise of the Subversive Multivector Threat, in Cybercrime and Espionage: An Analysis of Subversive Multi-Vector Threats 1 (Syngress 2011)..

threats (APTs). This implies that they do long-term harm by going unnoticed in a network for a long time[10].

Hackers can gather the data they need to initiate complex social engineering techniques or even produce original malware by persistently infecting the system. These multi-step attacks are extremely hard to identify and can persist for months or even years. Supply chain attacks are another important tactic used by these threat actors, who will target defense contractors and outside service providers that collaborate with government organizations. This method has the potential to destroy entire sectors and result in financial losses[11].

Potential targets must put advanced monitoring, threat intelligence, and intrusion detection systems in place because nation-state actors pose a threat to national security. In order to stop these external threat actors from endangering the public, quick incident reaction is essential.

- Cyber Criminals-An individual who commits crimes using computers and digital systems is known as a cybercriminal. Identity theft, phishing schemes, and credit card fraud are a few of the most prevalent cybercrimes. Although many hackers work alone, it's also typical for these threat actors to collaborate to do more harm.

Usually, money is the driving force behind cybercrime rather than politics or personal issues. They prey on both people and companies, stealing money through extortion, theft, and other financial fraud schemes. Cybercriminals will also steal intellectual property and private information, which they will

[10] Alexander Klimburg, The Whole of Nation in Cyberpower, in International Engagement on Cyber: Establishing International Norms & Improved Cybersecurity, 12 Geo. J. Int’l Aff. 171 (2011).

[11] Joab Kose, Cyber Warfare: An Era of Nation-State Actors and Global Corporate Espionage, ISSA J., Apr. 2021, at 12

then sell on the dark web to get money. Numerous cybercriminals target thousands of people with their campaigns, making them mass fraudsters.

- Hacktivists and Ideological actors- Hacktivists and ideological actors have quite different objectives than other threat actors, who usually have malevolent purposes. Hacktivists target particular companies to express their political or social views. Hacktivists target groups or individuals they believe to be at odds with their values. Although the goal of these attacks is to bring about societal change, this isn't always the result. For attention, they frequently target government figures or high-level executives. A hacktivist may initiate these ideological attacks for a variety of reasons. These include campaigns for human rights, environmental action, or criticism of a particular business or government[12].

- Insiders and Competitors- Insider risks, also known as internal threat actors, are individuals that work for a company and utilize technology to harm others. Current and past workers, independent contractors, and service providers may fall under this category. Competitors of the brand may potentially pose a threat by employing malevolent tactics to breach networks and steal intellectual property or business plans. Threat actors that engage in corporate espionage employ a range of common hacking methods, such as social engineering, ransomware campaigns, and taking advantage of security flaws. They might, however, also deploy insider recruitment techniques, snatching information from existing or former workers. In certain instances, they might exploit physical proximity to carry out their attacks, spying on people within the same network through Wi-Fi eavesdropping.[13]

[12] Nicolò Bussolati, The Rise of Non-State Actors in Cyberwarfare, in Cyber War: Law and Ethics for Virtual Conflicts 102 (Jens David Ohlin, Kevin Govern & Claire Finkelstein eds., Oxford Univ. Press 2015).

[13] Johan Sigholm, Non-State Actors in Cyberspace Operations, J. Mil. Stud., https://doi.org/10.1515/jms-2016-0184.

2. Existing International Laws and Regulations Governing Cyber Warfare

The Tallinn Manual, previously known as the Tallinn Manual on International Law Applicable to Cyber Warfare, is an academic, non-binding study of how international law applies to cyber conflicts and cyber warfare. This guidebook is one of the most reputable sources of laws connected to cyberwarfare and is extremely useful in defining and dealing with cyberwarfare rules. In one of the chapters, named 'Means and Methods of Cyberwarfare, some of the rules relating to those methods are-

➢ *RULE 44: "Booby Traps"*— The rule prohibits the use of cyber booby traps with specific objects listed in the law of armed conflict. This rule is based on the Mines Protocol and Amended Mines Protocol. Consider an email with a virus attachment, such as a kill switch, sent to a water treatment plant employee claiming to be from their physician. When the malware is activated, it suspends the purification process at the facility that serves both military and civilian users. This allows untreated water to enter the water supply used by soldiers. Illness is the intended outcome.

➢ *RULE 45: "Starvation"*— This rule prohibits using starvation to harm civilians during cyber warfare. In this manual "starvation" refers to intentionally depriving a civilian population of food and water in order to weaken or kill them.

➢ *RULE 46: "Belligerent Reprisals"*— Prohibits cyber operations against prisoners of war, interned civilians, those hors de combat, and medical personnel, facilities, vehicles, and equipment. Belligerent reprisals are activities that would violate armed conflict law if not in reaction to the enemy's transgressions. Reprisals should only be used to convince the opponent to follow the law. They differ from retribution, punishment, and retaliation by focussing solely on ensuring future obedience from the other side.

➢ *RULE 48: "Weapons Review"*— All states must ensure that the cyber means of warfare they acquire or utilise comply with the standards of armed conflict law that bind the state. The legality of a method of cyber warfare must be determined by referencing its regular expected use at the time of review. If a means or method of cyber warfare is being created for immediate operational use, the lawyer advising the commander is responsible for ensuring compliance with the State's international law duties.

The Oslo Manual on the Law of Armed Conflict is another important source in this area that also includes laws expressly for cyberwarfare. It has been stated therein that during the creation of the Oslo Manual, states had differing interpretations of terminology like "cyber means of warfare," "cyber methods of warfare," and "cyber-attacks." Despite differences of opinion, the Group of Experts agreed with the International Court of Justice's nuclear weapons Advisory Opinion that the LOAC principles apply to all types of conflict. This includes cyberwarfare. The rules are as follows:

❖ *Rule 20 (a):* For the purposes of this Manual, "cyber operations" are operations that employ capabilities aimed at achieving objectives in or through cyberspace

❖ *(b)* Cyber operations, when carried out as part of an armed conflict, are governed by applicable principles and rules of LOAC.

❖ *Rule 21:* With respect to an armed conflict, States bear responsibility for their cyber operations as well as other activities conducted in cyberspace that are attributable to them. Such responsibility includes actions by all persons belonging to the armed forces of the State.

❖ *Rule 22:* All those involved in the conduct of cyber operations, including attacks, are responsible for their respective roles and, commensurate with their involvement, have obligations to ensure that such operations are conducted in accordance with the applicable principles and rules of LOAC.

- ❖ *Rule 24:*_In cyber operations occurring during an armed conflict, the concept of attack applies to all acts of violence against the adversary, whether in offence or defence. The acts must be intended to cause—or must be reasonably expected to result in—death, injury, destruction or damage. These acts generally do not include those intended to cause only temporary loss of functionality.

- ❖ *Rule 27:* The concept of direct participation in hostilities applies to civilians, including civilian employees of State agencies, who conduct cyber operations in the context of an armed conflict.

- ❖ *Rule 28:* Cyber operations qualifying as direct participation in hostilities may include: (a) Any cyber activity designed or intended to directly cause death, injury, damage or destruction to an adverse party; (b) Cyber defence of military objectives against enemy attacks; (c) Contributing to targeting procedures, such as helping to identify or prioritize targets; (d) Engaging in planning specific cyber-attacks; or (e) Providing or relaying information of tactical relevance for the purpose of aiding in combat operations.

- ❖ *Rule 29:* In cyber operations constituting attacks, feasible precautions should be taken where necessary in order to avoid, or in any event minimize, destruction or damage to civilian objects, or death or injury to civilians.

- ❖ *Rule 30*: A Belligerent State should not conduct cyber operations that constitute attacks causing physical damage to or destruction of objects located in neutral territory, including neutral cyber infrastructure, unless the Neutral State is unable or unwilling to terminate an abuse of such objects or infrastructure by an adversary of the Belligerent State.

- ❖ *Rule 31:* Belligerent States must not launch attacks from cyber infrastructure located in neutral territory or under the exclusive control of Neutral States.[14]

2.1 Cyber Warfare in the Context of International Humanitarian Law

In the context of International Humanitarian Law, IHL consists of hundreds of rules; these are some of the rules that anyone conducting a cyber operation in the context of an armed conflict (including non-State armed organizations and civilian hackers) must be aware of and obey at the very least. Groups or collectives should guarantee that their members adhere to these boundaries. [15]

- *The Principle of Distinction*— According to Article 48 of the API, "the Parties to the conflict shall at all times distinguish between the civilian population and combatants, as well as between civilian objects and military objectives, and accordingly shall direct their operations only against military objectives." Because of the dual-use nature of cyberspace, the line between civilian and military cyber infrastructure is less pronounced, making it difficult to identify legitimate targets. The principle of distinction, which compels nations to distinguish between civilian and military troops and limit attacks to military objectives, requires cyber conflict sides to refrain from committing acts that might cause significant collateral damage. A cyberattack that attacks a military air traffic control system but only causes a troop transport to crash would be in accordance with the concept of distinction. Other cyberattacks, such

[14]Yoram Dinstein & Arne Willy Dahl, Oslo Manual on Select Topics of the Law of Armed Conflict: Rules and Commentary (Springer 2020), https://doi.org/10.1007/978-3-030-39169-0.)

[15] Tilman Rodenhäuser & Mauro Vignati, 8 Rules for "Civilian Hackers" During War, and 4 Obligations for States to Restrain Them, Int'l Comm. Red Cross: Humanitarian L. & Pol'y Blog (Oct. 4, 2023), https://blogs.icrc.org/law-and-policy/2023/10/04/8-rules-civilian-hackers-war-4-obligations-states-restrain-them/.

as those on civilian banking institutions, hospitals, museums, or places of worship, would clearly breach the principle of distinction.

- *The Principle of Proportionality-*— Since civilian deaths and/or destruction of civilian objects are unavoidable in combat, the principle of proportionality is one of the most contentious aspects of IHL. Under Article 51(5)(b), an attack is unlawful if it "may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated." Cyber-attacks present particular issues in terms of proportionality because of the nature of the harm they do. It can be difficult to determine whether an attack is proportional based on the relevant categories of "loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof," because the typical direct effects of cyber-attacks may be nonlethal or temporary but severe. A cyber-attack that successfully halts the delivery of information across the Internet may only disturb the public, but it might also have more serious effects. For example, it may prevent hospitals from communicating critical information, resulting in loss of life. Thus, attacks may alter the weight accorded to fleeting repercussions, forcing nations to confront greater ambiguity than they normally do when deciding on the legality of planned strikes.

- *The Principle of Precaution-*— It is a jus in bello principle that requires parties participating in cyber warfare to take precautions to minimise harm to people and civilian infrastructure. In the context of cyber warfare, the principle of precaution can be applied by doing a detailed risk assessment that considers the hazards and repercussions of cyber-attacks. This includes assessing the likelihood of collateral damage to civilian infrastructure, disruption of critical services, and harm to noncombatants.

The principle of proportionality also demands parties to assess the expected military advantage in their response to an adversary strike against the possible harm to civilians and civilian infrastructure. That is, the anticipated benefits of a cyber operation must outweigh any possibility of collateral damage. [16]

3. Forms and Methods of Cyber Warfare

While means of warfare relate to the weapons or equipment used in conflict, methods of warfare are tactics or plans employed to gain an advantage or weaken the opponent during military operations. For example, ruses are a legitimate and widely recognized tactic in armed engagements. Among many other strategies, Ruses include adopting camouflage, pretending lethargy or activity, and using decoys or dummy materials. Examples of illegal means of warfare include perfidy, the use of human shields, and the misuse of protected insignia[17].

A definitional foundation for the terms "means and methods of warfare" in the context of cyberspace is provided by the Tallinn Manual 2.0. Cyber weapons and related systems are included in the Manual's definition of "cyber means of warfare," as are any cyber devices, materials, instruments, mechanisms, equipment, or software that are used, intended to be utilized, or developed to be used in a cyberattack[18]. Cyber weapons are military tools that are used, created, or intended to harm or kill people or destroy or damage property. Lastly, according to the Tallinn Manual 2.0, "the cyber tactics,

[16] Chukwudumebi O. Joseph-Asoh, Nkechinyere Worluh-Okolie & Jojo Ebibode, The Rise of Cyberwarfare: The Applicability of International Humanitarian Law for the Protection of Civilians and Civilian Objects, 10 Int'l J.L. 1 (2024), https://www.lawjournals.org/archives/2024/vol10/issue2/10065.

[17] Geoffrey S. Corn et al., The Law of Armed Conflict: An Operational Approach 288 (2d ed. Wolters Kluwer 2019).
U.S. Dep't of the Army, The Commander's Handbook on the Law of Land Warfare, FM 6-27, MCTP 11-10C, at 2-1 (Aug. 2019).

[18] Michael N. Schmitt, ed., Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations 401 (Cambridge Univ. Press 2017).).

techniques, and procedures by which hostilities are conducted are the methods of cyber warfare." Cyber warfare techniques include hacking, phishing, distributed denial of service, and the employment of so-called honeypots and watering holes[19].

Multiple kinds of Cyber-attacks include:

- Espionage: Espionage in cyberwarfare aims to discover as much as possible about the adversary's informational, physical, and cybersecurity assets. These objectives are the same as those in conventional warfare. This may entail mapping target networks, breaching data sources, and then exfiltrating information through the use of social engineering, phishing, and server and network hacking.

- Cyber Sabotage: In a cyberwar, sabotage aims to undermine, disable, tamper with, or destroy a target's cybersecurity defenses, information services, and resources. A prime example of cyber sabotage is the Stuxnet malware, which was created by Israel and the United States to harm Iran's nuclear fuel processing capabilities.

- Cyber Psychological Warfare and Propaganda: Since the beginning of time, conventional warfare has employed psychological warfare, or PsyOps: Using disinformation campaigns, ransomware attacks, website takeovers and defacements, and distributed denial of service attacks to render websites and services inoperable, nation-states employ cyber PsyOps to wreak societal havoc in the digital age.

Some of the techniques used in Cyber Warfare include:

[19] Jeffrey T. Biller & Michael N. Schmitt, Classification of Cyber Capabilities and Operations as Weapons, Means, or Methods of Warfare, 95 Int'l L. Stud. 179 (2019), https://digital-commons.usnwc.edu/ils/vol95/iss1/6/.

- Distributed denial of service- DDoS attacks employ a variety of tactics to bombard a target with fictitious requests, which can interfere with or halt systems and operations and prevent civilians, security and military personnel, or research organizations from accessing websites. DDoS assaults are the most straightforward to launch out of all the many kinds of cyberattacks. U.S. and U.K. officials blamed Russia for a string of DDoS attacks that momentarily shut down the Ukrainian government and banks in the early days of Russia's invasion of Ukraine. Washington is trying to hold Russia accountable for its aggressive actions in cyberspace, U.S. deputy national security adviser Anne Neuberger told reporters at the White House[20].

- Ransomware- A ransomware assault is based on the deployment of malware that encrypts a computer's disk drives and demands payment, typically in cryptocurrency. These attacks, which are most frequently initiated through phishing, have emerged as one of the most popular methods for denial of service and extortion. According to the 2021 "Verizon Data Breach Investigations Report," ransomware accounted for 10% of all breaches, and 2,084 ransomware complaints were submitted between January and July 31, 2021, according to the FBI's Internet Crime Complaint Center. This is a 62 percent increase in ransomware year over year[21].

- Phishing, Spearphishing, and Whaling- Phishing attacks aim to compromise a victim's computer, cybersecurity, and network connections by targeting anyone who might click on a link in an email. Spearphishing

[20] Raphael Satter, US, UK: Russia Responsible for Cyberattack Against Ukrainian Banks, Reuters (Feb. 19, 2022), https://www.reuters.com/world/us-says-russia-was-responsible-cyberattack-against-ukrainian-banks-2022-02-18/.

[21] Verizon, Data Breach Investigations Report, https://www.verizon.com/business/resources/reports/dbir/ (last visited Apr. 7, 2025)..

is a more targeted attempt to obtain access to a firm by targeting individuals who work for that business or industry. Attackers may use drive-by downloads, dynamic URLs, and email spoofing to carry out a spearphishing assault and get beyond security measures[22].

High-level executives who have access to financial information or organizational data are the precise targets of whaling assaults. For instance, a C-level executive may contact an executive with financial approval authority requesting that they provide a sizable sum of money immediately to fulfill a vendor payment or other comparable commitment.

4. Notable Cyber Warfare Incidents and Their Outcomes

1. Stuxnet

Stuxnet was the term given to a highly complex digital malware that targeted and physically harmed Iran's clandestine nuclear program between 2007 and 2010, when computer security researchers revealed its existence. The malware targeted computer systems that manage physical infrastructure, such as centrifuges and gas valves.[23] Stuxnet took advantage of several previously undiscovered Windows zero-day vulnerabilities. That description should make it apparent that Stuxnet was part of a high-level sabotage campaign launched by nation-states against their opponents. It is now commonly acknowledged that Stuxnet was developed by the intelligence agencies of the United States and Israel. Stuxnet was first discovered by the infosec community in 2010, but development likely began in 2005. The United States

[22] William H. Boothby, Methods and Means of Cyber Warfare, 89 Int'l L. Stud. 387 (2013), https://digital-commons.usnwc.edu/ils/vol89/iss1/8/..

[23] Joshua Alvarez, Stuxnet: The World's First Cyber Weapon, Stanford Univ. Ctr. for Int'l Sec. & Cooperation (Feb. 3, 2015), https://cisac.fsi.stanford.edu/news/stuxnet.

and Israel wanted Stuxnet to derail, or at least postpone, Iran's nuclear weapons development program.[24]

The revelation of Stuxnet marked a turning point in how states perceive cyberthreats, making this hotspot analysis essential. Cyber Strategies differed significantly before and after Stuxnet. The Stuxnet worm is a computer malware that targets SCADA systems in industrial controls. It's unclear how the malware was created, but it undoubtedly took a significant amount of time, effort, and money to construct. Experts believe the worm's creation required a staff of 5-10 full-time programmers for at least six months. Stuxnet is larger than other worms, built in many programming languages, and includes encrypted components. The attacker used four zero-day vulnerabilities to infect machines, including an automatic procedure via USB devices, a link with shared printers, and two privilege escalation flaws.

Stuxnet appears to have targeted the Iranian nuclear reactor and uranium enrichment complex in Natanz. Stuxnet's targeting of devices in groups of 164 objects and Natanz's cascades of 164 centrifuges may not be coincidental. Iran employs inefficient and antiquated IR-1 centrifuges, a European model from the late 1960s and early 1970s. Centrifuges are fragile, and sudden changes in speed can damage or break them. The designers of Stuxnet were aware of and exploited this vulnerability. The Natanz nuclear reactor uses an air-gapped, closed computer network that is not connected to the Internet or other networks. Stuxnet most likely infiltrated the network using a detachable USB drive, implying that the worm's designers needed someone to transport it and infect the network.[25]

[24] Josh Fruhlinger, Stuxnet Explained: The First Known Cyberweapon, CSO Online (Aug. 31, 2022), https://www.csoonline.com/article/562691/stuxnet-explained-the-first-known-cyberweapon.html.

[25] Marie Baezner & Patrice Robin, Hotspot Analysis: Stuxnet, CSS Cyber Defense Project, Center for Security Studies (CSS), ETH Zürich (Oct. 2017), https://doi.org/10.3929/ethz-b-000200661.

On a domestic political level, the cyberattack harmed the Iranian regime since the Iranian authorities were unable to safeguard their nuclear assets from international cyberattack. The Iranian leadership appeared undecided on how to officially respond to reports that a computer worm may have compromised its nuclear facilities. In September 2010, the Iranian authorities first downplayed the impact of the attack in their discourse, most likely to avoid blame from the Iranian people, by claiming that only personal computers without connections to the Bushehr nuclear facility were infected and blaming the West and NATO. Two months later, they confessed that the worm had been present at their nuclear plants for over a year. However, they did not remain idle and worked tirelessly to contain and eradicate the infection, as well as identify the attackers. At the same time, Iranian officials did not retaliate against the cyberattacks since the culprits' identities were unknown or uncertain, and no precedence existed for how a state should respond to such an attack. This delay made the Iranian government appear weak and easier to target. The most significant impact of Stuxnet on society was likely a sense of insecurity, as an incursion into a private domain is never taken lightly, and it can thus be believed that Iranians felt betrayed by the country's insufficient cybersecurity measures and its weak attitude in regard to the culprits. [26]

2. WannaCry

The WannaCry ransomware attack, which happened in May 2017, is regarded as one of the most major and damaging cyberattacks in recent memory. It emphasises organisations worldwide the necessity of cybersecurity awareness and effective defence tactics. WannaCry was a type of ransomware that rapidly spread throughout the world, infecting over 200,000 systems in more

[26] Marie Baezner & Patrice Robin, Hotspot Analysis: Stuxnet, CSS Cyber Defense Project, Center for Security Studies (CSS), ETH Zürich (Oct. 2017), https://doi.org/10.3929/ethz-b-000200661.

than 150 countries. Unlike Locky, which spreads by phishing, the malware exploited a vulnerability in Microsoft Windows known as EternalBlue. The WannaCry ransomware attack had a devastating worldwide impact. It shut down businesses, delayed medical treatments, and resulted in severe financial losses. The total projected damage ranged between hundreds of millions and several billion dollars. Critical infrastructure, such as healthcare and transport, faced significant hurdles as they worked to restore service. The event emphasised the vulnerability of organisations using obsolete software and the potential implications of failing to prioritise cybersecurity measures.[27]

The WannaCry ransomware software operates in a straightforward manner and is not regarded exceptionally complicated or innovative. It comes on the infected machine as a dropper, which is a self-contained program that extracts the other application components placed within it. WannaCry spreads through a weakness in Microsoft Windows' implementation of the Server Message Block (SMB) protocol. The SMB protocol allows network nodes to interact, and an unpatched version of Microsoft's implementation might be tricked into executing arbitrary code via specially crafted messages, a vulnerability known as EternalBlue. The fact that this relatively basic malware was distributed via EternalBlue is arguably intriguing than the ransomware itself. It is thought that the United States National Security Agency found this weakness and, instead of reporting it to the information security community, created the EternalBlue malware to exploit it. This exploit was later stolen by the Shadow Brokers, who released it obfuscated in a purportedly political Medium article on April 8, 2017. Microsoft detected the weakness a month ago and published a fix, but many computers remained unpatched and unprotected, and WannaCry, assisted by EternalBlue, spread swiftly on May 12. Following the

[27] Lïa Desmousseaux de Givré, WannaCry Ransomware Attack: A Case Study, Arsen (Oct. 19, 2024), https://arsen.co/en/blog/wannacry-ran

outbreak, Microsoft chastised the US government for not disclosing the existence of the vulnerability earlier.[28]

The attack caused significant disruption to various organisations throughout the world, including the UK's National Health Service (NHS), which cancelled numerous appointments and operations, Spanish telecoms major Telefónica, and many other companies and institutions. Damages have been estimated at around $4 billion. The UK National Health Service (NHS) was one of the most seriously compromised organisations. The hack disrupted services, cancelling roughly 19,000 medical appointments and compromising multiple systems. The NHS's direct costs were approximately £92 million ($120 million), with an additional £72 million ($94 million) spent on IT improvements and cybersecurity concerns.[29]

3. Russian Attacks on Ukraine

On December 23, 2015, and December 17, 2016, online agents known as Sandworm, who are affiliated with Russia, targeted the Ukrainian power grid by deactivating the substations responsible for supplying electricity to communities. While Ukraine is an exception due to its usage of obsolete Soviet equipment, these cyber-attacks against Ukraine demonstrate the budding field of cyber warfare and the necessity for states to defend their power grids.[30]

On December 23, 2015, three Ukrainian power distribution firms' control centres were accessed remotely. Malicious actors gained access of the

[28] Lïa Desmousseaux de Givré, WannaCry Ransomware Attack: A Case Study, Arsen (Oct. 19, 2024), https://arsen.co/en/blog/wannacry-ransomware.

[29] Pete Barnum, WannaCry Ransomware Attack (2017) – Technical, Financial, and Legal Analysis, Inedo Security, https://security.inedo.com/library/incidents/wannacry-2017 (last visited Apr. 8, 2025).

[30] Miles Pollard, A Case Study of Russian Cyber-Attacks on the Ukrainian Power Grid: Implications and Best Practices for the United States, 16 Pepp. Pol'y Rev. 1 (2024), https://digitalcommons.pepperdine.edu/ppr/vol16/iss1/1/.

facilities' SCADA systems and opened breakers at approximately 30 distribution substations in the capital of Kiev and the western Ivano-Frankivsk area, causing more than 200,000 users to lose electricity. Over a year later, on December 17, 2016, a single transmission substation in northern Kiev went dark. These acts of sabotage occurred following a political upheaval in Kiev, the annexation of Crimea, and military conflicts in eastern Donetsk and Luhansk areas. Governments and cybersecurity firms have blamed the hacks on Russian groups suspected of having ties to the Russian government, though the link is unclear.[31]

This attack could have been a reaction to a pro-Ukrainian separatist attack in Crimea. Since the separatists are accused of cutting off power to two million people in Crimea and the vital naval base at Sevastopol, this attack might have served as a warning to Ukraine that Sandworm is capable of much more and has been operating within vital utilities for months without being detected. The power grid hacks in Ukraine have provided ample proof of the growing cyber threat to industrial control systems. Electricity loss in the modern economy has disastrous effects for both production and security, regardless of whether it is generated, transmitted, or distributed. During outages, not only do factories and offices cease operations, resulting in lost worker productivity, but water treatment facilities and hospitals are unable to provide for inhabitants, potentially leading to widespread health outbreaks.[32]

[31] Donghui Park & Michael Walstrom, Cyberattack on Critical Infrastructure: Russia and the Ukrainian Power Grid Attacks, Jackson Sch. Int'l Stud., Univ. of Wash. (Dec. 2016), https://jsis.washington.edu/news/cyberattack-critical-infrastructure-russia-ukrainian-power-grid-attacks/.

[32] Miles Pollard, A Case Study of Russian Cyber-Attacks on the Ukrainian Power Grid: Implications and Best Practices for the United States Grid, 13 Pepp. Pol'y Rev. 1 (2020).

5. Necessary Legal Reforms in Cyber Warfare under International Law

The rapid advancement of cyber capabilities has outpaced the development of international legal frameworks governing cyber warfare. While existing international laws, including the UN Charter, the Geneva Conventions, and customary international law, provide some guidance, they are often inadequate to address the unique challenges posed by cyber warfare. The increasing frequency of cyberattacks by both state and non-state actors necessitates urgent legal reforms to ensure accountability, deterrence, and stability in cyberspace.

One of the most pressing reforms is the explicit recognition of cyber warfare within international law. The UN Charter prohibits the use of force except in self-defense or with Security Council authorization. However, it remains unclear whether cyberattacks that cause significant harm, such as disabling critical infrastructure or financial systems, qualify as an armed attack under Article 51[33]. A revised framework must establish clear thresholds for determining when a cyber operation constitutes an act of war, warranting a proportional response.

Attribution is another critical challenge in cyber warfare. Unlike traditional military attacks, cyberattacks often involve layers of obfuscation, making it difficult to identify perpetrators with certainty. International law should establish mechanisms for cooperative attribution efforts, such as an independent body that verifies cyber incidents and provides credible attribution reports. Strengthening state responsibility by holding nations accountable for cyber activities originating from their territory, even if conducted by non-state actors, is another necessary reform[34].

[33] Michael N. Schmitt, Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations r. 103 cmt. (Cambridge Univ. Press 2017)..

[34] Jeffrey T. Biller & Michael N. Schmitt, Classification of Cyber Capabilities and Operations as Weapons, Means, or Methods of Warfare, 95 Int'l L. Stud. 179, 219 (2019).

Additionally, the application of the principle of distinction in cyber warfare requires further refinement. Under international humanitarian law, combatants must distinguish between military and civilian targets, yet cyberattacks frequently blur these lines. A legal framework must clarify what constitutes a legitimate military cyber target and establish safeguards to minimize collateral damage to civilian infrastructure, such as hospitals, water systems, and communication networks.Another crucial area for reform is the regulation of offensive cyber capabilities[35]. While states have the sovereign right to develop cyber defenses, there is little regulation regarding the use of cyber weapons. An international agreement akin to arms control treaties could establish limitations on the deployment of destructive cyber tools and encourage transparency measures, such as the notification of newly developed cyber weapons to an oversight body.

Furthermore, the protection of critical infrastructure against cyber warfare requires enhanced legal mechanisms. Cyberattacks on power grids, financial systems, and election infrastructure pose severe threats to national security and democratic stability. International law should mandate cooperative measures between states to prevent attacks on such critical infrastructure and establish consequences for those who target them[36].

Finally, there is a need for stronger enforcement mechanisms in cyber law. Currently, responses to cyberattacks often rely on diplomatic or economic sanctions, which may not always be effective. A specialized international cyber tribunal could be established to adjudicate cyber disputes and impose penalties for violations of cyber warfare laws. This would provide a legal

[35] Nicolò Bussolati, The Rise of Non-State Actors in Cyberwarfare, in Cyber War: Law and Ethics for Virtual Conflicts 102, 102–26 (Jens David Ohlin, Kevin Govern & Claire Finkelstein eds., Oxford Univ. Press 2015).

[36] Raphael Satter, US, UK: Russia Responsible for Cyberattack Against Ukrainian Banks, Reuters (Feb. 19, 2022, 12:58 PM), https://www.reuters.com/world/us-says-russia-was-responsible-cyberattack-against-ukrainian-banks-2022-02-18/.

avenue for addressing cyber conflicts rather than resorting to retaliatory cyber operations.

6. Conclusion

In an era where technology permeates every aspect of society, cyber warfare has emerged as a significant challenge to global security and legal frameworks. The intersection of cyber warfare and International Humanitarian Law (IHL) highlights the complexities of applying traditional wartime regulations to digital battlegrounds. While existing IHL principles, such as distinction, proportionality, and necessity, provide a foundational framework, they often struggle to address the unique characteristics of cyber conflicts, including attribution difficulties, civilian infrastructure vulnerabilities, and the blurred lines between military and non-military actors. the role of non-state actors, including hacktivists, private corporations, and terrorist organizations, further complicates the regulatory landscape. Their involvement raises critical questions about accountability and the applicability of IHL principles in conflicts where state actors are not the sole participants. To address this, international legal bodies must explore innovative legal interpretations and enforcement mechanisms, ensuring that cyber warfare remains governed by principles that safeguard civilian populations and critical infrastructure.

Ultimately, while cyber warfare presents new and evolving challenges, the core tenets of IHL must adapt to ensure that armed conflicts, regardless of their domain, adhere to humanitarian principles. The international community must act decisively to bridge legal gaps, enhance cooperation, and reinforce ethical norms to prevent cyber operations from becoming unchecked instruments of destruction. Without proactive measures, the world risks facing an era where warfare is not only digital but also lawless, endangering global stability and human rights.

THE METAVERSE: A TRANSFORMATIVE DIGITAL PARADIGM RESHAPING HUMAN INTERACTION AND SOCIETY

Riya Sharma and Surbhi Bharti[1]

***Abstract**:*

The metaverse, being a set of changes in the way people interact, perform transactions, and govern themselves in digital space, is studied as a complex evolving ecosystem created by the confluence of blockchain technology, AI, AR/VR, and decentralized networks. The study applies a mixed-methods approach by analysing quantitatively the trends in user engagement and digital assets and conducting qualitative case studies on platforms such as Decentraland, The Sandbox, NVIDIA Omniverse, and Meta Horizon Workrooms. It is a highly critical examination of the technologies forming the underlying infrastructure, economic opportunities, social interactions, and governance modalities emerging within the metaverse.

According to the findings, blockchain platforms empower decentralized ownership and governance via organizations called DAOs. Meanwhile, AI and immersives promote the user experience with ethical concerns around privacy, bias, and mental health. Economically, the metaverse channels money from various streams through DeFi, virtual real estate, and branded digital commerce while regulatory issues, accessibility, and market fluctuations still tussle for attention. Glimpses of the new digital identity, new norms in interaction with communities, and inclusion also increase with risk in harassment, data exploitation, and psychological targeting. These, in turn,

[1] 4th Year B.B.A. LL.B. (Hons.) students.

threaten decentralization and democratic participation ideals with an underdeveloped legal framework and centralized power concentration among wealthy users.

This study concludes that while metaverse holds extreme transformative possibilities for industries and social systems, it must be steered by ethical design principles, economically inclusive models, and globally coordinated governance structures. It further recommends that international regulatory standards be developed, privacy-by-design frameworks be embedded, digital equity be promoted, and interdisciplinarity be fostered so that the metaverse can evolve into sustainable, inclusive, and legally coherent space.

Keywords: Metaverse, Blockchain Technology, Decentralized Autonomous Organizations (DAOs), Artificial Intelligence, Augmented Reality (AR) & Virtual Reality(VR)

Introduction

The digital landscape is undergoing a monumental transformation With the conception of the metaverse, there will be a major transformation in the digital landscape: a series of interconnected virtual environments where users can interact, collaborate, and carry out transactions in immersive digital settings. "Metaverse" or "computer-simulated universe" was a term possibly first introduced by Neal Stephenson in his 1992 novel titled Snow Crash. In the present evolution, it comprises the entire digital ecosystem supported by a combination of blockchain technology and AI, AR, VR, and decentralized networks. (Stephenson, 1992)[5]

The metaverse is essentially a network of interconnected digital environments where one can perform various human-like activities-gaming, socializing, working, learning, conducting business, and more. Today, the blend of many technologies is elevating the metaverse from a conceptual idea into a technological reality, thus attracting investments by the largest corporations, financial institutions, and governments. (Radanliev, L., et al., 2023)[2]

Major technology companies like Meta (formerly Facebook), Microsoft, and NVIDIA leading the charge, major tech companies are pouring billions into the construction of immersive platforms and virtual worlds that promise to redefine the way people associate with each other on the Internet. On the other hand, blockchain-based platforms such as Decentraland or The Sandbox have also shown the possibilities for decentralized governance, in which case the user can own digital assets, vote on platform policies, and partake in decision-making processes. (Zallio, M., & Clarkson, P. J., 2023)[1]

Whereas the very technology that forms the foundation of the metaverse offers clear opportunities, equally pressing ethical, societal, and governance questions emerge. These issues about privacy, about digital identity and data security, about mental wellness, about corporate monopolization need to be addressed responsibly. (Garon, J. M., 2022)[3]

The present paper intends to serve as a broad overview of the technological underpinnings, economic concerns, social and ethical considerations, and governance pertinent to the metaverse. The study, through detailed case studies and theoretical analysis, hopes to continue contributing to the stagnating conversation on how the metaverse can be developed in a manner that promotes inclusivity, ethical integrity, and sustainable progress. (Zhang, X., Chen, H., & Liu, Y., 2023)[4]

I. Technological Foundation

This research paper makes a very comprehensive and well-integrated study of the technology foundations of the metaverse. It looks carefully into the merging of Blockchain Technology, Artificial Intelligence, AR/VR, and

Decentralized Networks in the realization of the immersive, interactive, and decentralized fabric of the metaverse.

Blockchain Technology

Blockchain technology forms the basis for decentralized ownership and transaction systems within the metaverse. It provides a transparent, immutable ledger that protects users in trading digital assets such as NFTs, virtual real estate, and digital art. The decentralized view provided by a blockchain eliminates the need for intermediaries, thus allowing users to basically trade on a peer-to-peer basis and, therefore, perceiving a higher degree of autonomy (Zallio & Clarkson, 2023)[1].

Case Study: *Dectraland & The Sandbox*

Decentraland and The Sandbox are two major blockchain ecosystems that the stage for decentralized governance and digital asset ownership are set upon. In Decentraland, users are free to buy, sell, and create on their land using MANA, the native cryptocurrency. On the other hand, governance is treated as a major factor through their DAO-based system, permitting users to vote on matters concerning changes to policies so that the platform is built within consensus.

The Sandbox operates on a somewhat similar line but has somewhat more emphasis on user content and gaming. In 2021, Republic Realm, a digital real estate investment firm, purchased a plot of land in The Sandbox for about

$4.3 million-the one that is among the most significant in the metaverse. (Zallio & Clarkson, 2023)[1]. This transaction underlines the extent to which virtual land is now perceived as a valuable economic resource and demonstrates how blockchain can be used for ownership of digital assets.

2. Artificial Intelligence (AI)

Artificial Intelligence acts as the crucial element in the metaverse; it is what powers intelligent avatars, natural language processing, automated content generation, and recommendation systems. AI-based avatars learn from their interactions, evolve with user preferences, and provide conversational experiences very much akin to a human. (Garon, 2022)[3].

Case Study: *NVIDIA omniverse*

NVIDIA's Omniverse is a next-generation AI-powered platform giving collaborators a new way to conjoin virtual environments. Developers can, through the Omniverse, construct the digital worlds and systems for hypotheses of intricate physical systems. Omniverse AI avatars are the simulated entities meant to interact naturally with users to somewhat heighten the immersive feel (Garon, 2022)[3].

3. Augmented Reality (AR) & Virtual Reality (VR)

AR and VR technologies constitute the immense landscape of the metaverse in which immersive worlds are possible. Whereby a VR gives a user a fully immersive experience, AR enhances real-world interactions by overlaying digital information on the physical world.

Case Study: *Meta's Horizon Workrooms*

In the virtual world, Meta's Horizon Workrooms serve as a tool for collaboration, wherein users interact in the immersive meeting spaces through their VR headsets. It intends to elevate remote work experiences by providing high-fidelity virtual environments where colleagues gather to collaborate, share ideas, and engage in interactive discussions. (Zhang, Chen & Liu, 2023)[4].

4. Decentralized Networks

Decentralized networks promote data privacy, security, and user empowerment by enabling peer-to-peer interactions without the need for central authorities. They form the foundation of platforms that operate on blockchain technology and ensure that users have control over their digital identities and assets. (Radanliev et al., 2023)[2].

Case Study: *Decentralized Autonomous Organisations(DAOs)*

DAOs provide a governance model that allows users to participate in decision-making processes through voting mechanisms. Decentraland's DAO, for example, offers token holders the ability to propose and vote on changes to the platform's policies, promoting a more democratic and decentralized governance structure.

Strengths

This discussion on blockchain technology is particularly interesting in how it explains decentralized ownership, peer-to-peer security in transactions, and governance through smart contracts. The paper offers the case studies of Decentraland and The Sandbox as illustrative examples of how users can own virtual land, transact in native cryptocurrencies (e.g., MANA), and govern the platforms together via Decentralized Autonomous Organizations (DAOs). These instances show how blockchain technology fosters economic activity within virtual worlds while changing the very nature of user empowerment and platform governance (Zallio & Clarkson, 2023; Radanliev et al., 2023)[1,2]. Similarly, another convincing pursuit is that of AI. The paper states rightly how AI is essential in creating and empowering adaptive avatars, natural language processing, and intelligent environmental responsiveness. Including the case study of NVIDIA Omniverse demonstrates the capability of AI to deliver highly realistic virtual worlds, support real-time collaboration, and enable sophisticated simulation functionalities (Garon, 2022)[3]. AI technologies inside the metaverse further stimulate personalization and user engagement, converting passive virtual spaces into dynamic, evolving ecosystems.

The treatment of AR and VR technologies within the paper is relevant and timely. It denotes the immersive capability of such technology to create an environment of "presence" and emotional realism. The case of Meta’s Horizon Workrooms is especially useful in highlighting how VR can simulate office settings and enact a level of interaction that is not entirely feasible through traditional videoconferencing (Zhang, Chen & Liu, 2023)[4]. The effectiveness of the metaverse in the context of remote working and education is issued, with another various other important issues for communication and productivity brought into view.

On the other side of decentralized networks, indeed, the paper exhibits the essence of peer-to-peer connectivity as the very foundation of users' sovereignty and data autonomy. The mention of DAOs, and especially that of Decentraland’s DAO, reveals an instance of a decentralized governance framework whereby users maintain the power to propose and vote upon changes affecting platform policies at large. This veers considerably from a centralized form of control and represents a somewhat matured view of digital democracy (Radanliev et al., 2023)[2].

Real-world case studies such as those of NVIDIA Omniverse, Meta's Horizon Workrooms, and Decentraland's DAO forge an analytical path and empirical support, giving the theoretical discussion a grounded, practical dimension. Each case study is effectively selected and contextually relevant, demonstrating how technological innovations are being translated into functional metaverse applications.).

II. Economic Implications

The paper deliberates upon the newly emerging economic landscape within the metaverse in a comprehensive and forward-looking manner. It first discusses how blockchain technology, Defi, NFTs, and virtual real estate are transfiguring conventional economic models and causing an industrial sector to arise whose nature and configuration are co-temporal with the Digital or

Web 3.0 economy. Drawing real-world examples and instituting agencies therein gives further validity and timeliness to the discourse.

1. **Cryptocurrencies & Decentralized Finance (DeFi)**

While the metaverse propagated and incubated cryptocurrencies and DeFi systems, the presence of these financial mechanisms allows the user engage in peer-to-peer transactions, trade digital assets, and participate in a decentralized economic system.

Case Study: *JPMorgan Chase in Decentraland*

In February 2022, JPMorgan Chase entered into the metaverse as the first big bank by opening a virtual lounge in Decentraland called Onyx. From here, the bank started to explore the possibilities of decentralized financial mechanisms and the development of financial products for the digital environment. In essence, JPMorgan Chase is trying to give financial services within the metaverse to streamline the gap between traditional banking and decentralized financial models. (Zallio & Clarkson, 2023)[1].

Case Study : *Etherum-Based DeFi Platforms*

Since blockchain is the principal blockchain network, it hosts many dApps that provide financial services like lending, borrowing, and trading. Aave and Compound have become trustworthy DeFi protocols where users can transact with each other without intermediaries. The amalgamation of such systems will enable better accessibility and foster financial inclusion into the metaverse. (Zallio & Clarkson, 2023)[1].

2. **Virtual Real Estate and Digital Commerce**

The idea of virtual real estate has indeed acquired some prominence whereby users purchase, develop, or monetize digital patches of land within different metaverse platforms.

Case Study: *Republic realm's Record-Breaking Purchase*

In 2021, Republic Realm, a digital real estate investment company, sold virtual land in the Sandbox for nearly $4.3 million. This transaction represents one of the largest sales in the virtual real estate market, showcasing the appreciation of digital assets within the metaverse. Republic Realm is working on drawing in users by creating immersive experiences, advertisements, and interactive environments. (Zallio and Clarkson, 2023)[1].

Case Study: *Nike's Nikeland on Roblox*

Nike has set up its digital slot called Nikeland on the Roblox platform so that users may indulge in sports games, perform challenges, and buy virtual Nike goods for their avatars. Nikeland takes a step forward in digital marketing and branding by offering users a marketing interface and interactive experiences to build customer engagement and loyalty. (Radanliev et al., 2023)[2].

Strengths:

The paper expertly arranges the rise of cryptocurrencies and DeFi mechanisms into history while enabling peer-to-peer transactions and decentralized banking processes in virtual worlds. By taking the examples of Ethereum-derived applications Aave and Compound, the study shows that smart-contract-driven ecosystems provide power to users for lending, borrowing, and trading digital assets without the intervention of traditional intermediaries (Zallio & Clarkson, 2023)[1]. This demonstrates a fundamental paradigm shift from finance being under control to a finance in which trust is based upon code, transparency, and cryptographic consensus.

The case study of JPMorgan Chase entering Decentraland offers a much-needed institutional anchor, thereby crystallizing the growing legitimacy of metaverse economics. By creating a virtual lounge called "Onyx," JPMorgan gave the definitive signal of how the potentially revolutionary financial services industry could be integrated with decentralized platforms, thus bridging legacy finance with Web3 innovation (Zallio & Clarkson, 2023)[1]. This reflects how various big financial houses have developed an interest in harvesting the value being created in immersive digital environments.

The real estate and digital commerce concepts are particularly insightful. The record-breaking land acquisition in The Sandbox by Republic Realm for a price of nearly $4.3 million depicts in gleaming colors the speculative and investment potential of buying virtual properties (Zallio & Clarkson, 2023)[1]. Furthermore, the case of Nike's Nikeland on Roblox creates a new dimension of interactive branding where companies develop immersive, gamified experiences to foster customer loyalty and monetize virtual products (Radanliev et al., 2023)[2]. These examples essentially point out that the metaverse is much more than a space for entertainment or social interaction-have become an economic universe.

The discussion then recognizes how NFTs are already becoming an essential part of the digital economy by giving verifiable ownership of digital art, fashion, and collectibles. NFT marketplaces embedded within metaverse platforms stimulate user creativity and entrepreneurship, thereby democratizing digital commerce and ownership.

Social Implications & Ethical Concerns

The paper indicates a thoughtful crate and multidimensional analysis of the social transformation and ethical dilemmas posed by the metaverse. As the digital environment increasingly comes in between human interactions, questions of digital identity, privacy, mental health, and inclusiveness position themselves as the putative kirijuitan and social recognition of metaverse.

3. Digital Identity & Privacy

The Forming of Digital Identities in the Metaverse allows novel ways of interaction and collaboration between users. With this ability, however, there arise concerns on privacy, especially cases where one's information is shared without the proper safeguards.

Case Study: *Meta's Horizon World's*

It has been said that Meta's Horizon Worlds faces public critique over privacy issues and unsatisfactory containment of harassment episodes. As a solution to cure the safety of users, Meta put in place the notion of "personal boundaries," thereby establishing a narrow approach to achieve this end. Nevertheless, in the eye of the critic, this does not go far enough, as the larger issue of data privacy and exploitation remains unaddressed. (Garon, 2022)[3].

4. **Mental Health & Well-Being**

Prolonged exposure to virtual environments may have detrimental effects on mental health, particularly if users become overly immersed in digital spaces at the expense of real-world interactions.

Case study: *Psychological impact of Metaverse Platforms*

Research by Zhang et al. (2023) highlights the risks associated with social isolation, addiction, and the erosion of real-world relationships resulting from prolonged immersion in the metaverse. Although the metaverse offers enhanced social presence and engagement, it also poses potential dangers to users' psychological well-being. (Zhang, Chen & Liu, 2023)[4].

Strengths:

The article aptly traces the digital identity from the prism of opportunities and risks. On one hand, the metaverse provides an enhanced ability to create and embody an identity that transcends physical or social limitations, thus enabling an increase in self-expression, creativity, and inclusiveness. On the other hand, this fluidity of identity raises profound concerns of identity theft, misconduct under the mask of anonymity, and user data manipulation.

Incidents of harassment and data misuse within Meta's Horizon Worlds are well cited as instances proving that the current limits within platform safeguards are insufficient. The introduction of such mitigations as "personal boundary" settings on Meta's part came only as a reaction, which somehow only brushed the surface of more profound issues concerning user safety and consent (Garon, 2022)[3].

An essential consideration is made for mental health and well-being issues. While the immersive aspect may enhance digital presence, it can also give rise to social isolation, addiction, and blurring the lines between the virtual and real. The referred research by (Zhang, Chen, and Liu 2023)[4] strengthens this concern, indicating that extended exposure to the metaverse environment might lead deteriorate real-world relationships and increase dependency on virtual validation. Although platforms offer avenues for virtual companionship and therapeutic interactions, they also risk replacing authentic human engagement, particularly among vulnerable populations.

Additionally, the paper insightfully flags ethical concerns about accessibility and inclusivity. It acknowledges the risk of a growing digital divide, where marginalized communities—due to lack of technological access, digital literacy, or economic capacity—may be excluded from participating meaningfully in metaverse experiences. Such exclusion not only reproduces real-world inequalities in virtual spaces but also poses challenges to the democratic ethos of the metaverse (Zallio & Clarkson, 2023)[1].

Privacy on data generated for surveillance capitalism is hence timely and necessary for consideration. With AI-driven systems tracking and recording each gesture of users, minute details about their behavioural patterns, facial expressions, voice and even motion within metaverse platforms, fears of the commodification of data, profiling and tracking without consent heighten; such threats get amplified when owners of these platforms begin to exploit such data for their most effective targeted advertising or behavioural manipulation without a regulation in place to check their blatant acts (Garon, 2022)[3]. The paper rightly points out that ethical frameworks, privacy-preserving technologies, and legislative safeguards must evolve alongside these platforms.

Governance Structures

The governance of the metaverse is one of its most complex and underdeveloped dimensions. The research paper demonstrates a clear understanding of the legal, structural, and normative challenges involved in regulating decentralized digital environments. It evaluates the promise of Decentralized Autonomous Organizations (DAOs) and the limitations of existing legal frameworks in addressing issues such as accountability, power concentration, and jurisdictional ambiguity.

1. Decentralized Governance Models

Decentralized Autonomous Organizations (DAOs) offer a promising approach to governance by allowing users to participate in decision-making processes through smart contracts. However, this model presents challenges related to power concentration and legal recognition.

Case Study: *Decentraland's DAO*

Decentraland's governance structure is based on a DAO where token holders can propose and vote on changes to the platform's policies. While this model promotes user participation, concerns about the disproportionate influence of wealthy participants remain prevalent. (Garon, 2022)[3].

Strengths:

The paper's exploration of decentralized governance models is timely and insightful. The paper uses Decentraland DAO's governance structure as a showcase to demonstrate how token voting truly gives end users the ability to decide on platform policies. This participatory model is an alternative to the traditional hierarchical governance system; whereas the metaverse leans towards decentralization and user sovereignty (Garon, 2022)[3]. The discussion wisely points out that these models endorse democracy by dispersing the decision-making power across a wide user-base, as opposed to a singular corporation or governing body.

Additionally, DAOs are not free from power asymmetries while being innovative. The power of rich token holders in the governance processes might compromise the ideals of democracy and bring about fears of plutocratic domination. This observation is crucial, reflecting the growth in maturity of how an economic capital finds its way into political capital even in decentralized ecosystems (Garon, 2022)[3]. The analysis further shows that decentralization does not necessarily mean justice or equality; hence it deserves continuous scrutiny.

Going back, the paper also stresses how a legal paradigm for metaverse governance is sorely lacking. Since the metaverse is trans-jurisdictional and its underlying technologies new, traditional regulatory mechanisms geared towards physical or Web2 contexts find it very hard to adapt. Intellectual property disputes, data protection obligations, taxation, and user accountability are all issues that remain poorly defined. The absence of any clear lines of authority and jurisdiction in virtual spaces comes as a blessing to defaulters who can now exploit legal loopholes and lay traps for unwary stakeholders (Garon, 2022)[3].

Findings

The research conducted offers amultidimensional understanding of the metaverse, being a big transformative digital ecosystem encompassing aspects of technological innovation, potential economic opportunities, social restructuring, and legal reforms. The findings are thus worked under four main heads: the technological, economic, social, and governance vectors, so as to give a complete synthesis of the opportunities and negative aspects that the advancements of virtual worlds pose.

5. Technological Findings

- The metaverse is basically where all the aforementioned brilliant technologies, including blockchain, AI, AR/VR, and decentralized networks, come together. They work synergetically in keeping digital spaces alive and computationally immersed with experiences.
- With blockchain platforms like Decentraland and The Sandbox, one can easily partake in decentralized ownership and transparent asset acquisition. Nevertheless, scalability issues prevail due to interoperability concerns, transaction speeds, and energy consumption (Zallio & Clarkson, 2023)[1].
- This further gets complicated due to AI integration, where user experiences are enhanced through intelligent avatars and content personalization in real-time on the one hand, and privacy, data sovereignty, and algorithmic bias get questioned on the other (Garon 2022)[3]. AR and VR hence contribute centrally toward the immersiveness, but with costs and access limitations, the underprivileged community might find barriers inhibiting full participation in the metaverse (Zhang, Chen & Liu 2023)[4].
-

6. Economic Findings

- The metaverse is a representation of a confluence of advanced technologies, such as blockchains, AI, AR/VR, and decentralized networks. These work synergistically to foster persistent, immersive, and interactive digital spaces.
- Blockchain platforms such as Decentraland and The Sandbox have indeed supported decentralized ownership and transparent transactions of assets. Nevertheless, issues regarding interoperability, transaction speeds, and energy consumption stand as hurdles for scalability (Zallio & Clarkson, 2023)[1].

- AI integration enhances user experiences through intelligent avatars and real-time content personalization but simultaneously raises concerns regarding privacy, data sovereignty, and algorithmic bias (Garon, 2022)[3]. Similarly, AR and VR technologies play a central role in immersion but are constrained by cost barriers and accessibility limitations, potentially excluding underprivileged communities from participating fully in the metaverse (Zhang, Chen & Liu, 2023)[4].

7. Social Findings

- The social life and interaction in the metaverse transcend the boundaries of their creation, collaboration, and places of creative expression. The ethical challenges are of serious magnitude. There are growing concerns about digital identities, data privacy, harassment, and mental health risks without the backing of strong safety regulations.
- Incidents of privacy violation and virtual harassment on platforms such as Meta's Horizon Worlds called for the installation of rudimentary interactivity controls such as personal boundaries, but this remained a very reactive and rather limited approach (Garon, 2022)[3].
- Studies on user behavior in an immersive environment suggest social isolation, addiction, and depersonalization as possible psychological side effects, especially for young users (Zhang, Chen & Liu, 2023)[4]. The digital divide also continues to hamper marginalized groups from fully participating, thus reinforcing existing social inequalities enacted through new digital tools (Zallio & Clarkson, 2023)[1].

8. Governance Findings

- The metaverse represents a basic challenge to our usual models of law and governance. The rise of Decentralized Autonomous Organizations indicates a hopeful shift toward user-led governance. The Decentraland DAO is a functioning example where token holders submit proposals and vote on them for platform policies, thus demonstrating the possibility of democratic participation in virtual spaces (Garon, 2022)[3].

- Yet, these cases reveal that power often tends to be concentrated: for example, where wealthier participants dominate decision-making. The absence of a full set of legal requirements, especially on jurisdiction, intellectual property, data protection, and moderation rights, acts as a roadblock toward able regulation and enforcement within digital spaces (Garon, 2022; Zallio & Clarkson, 2023)[3,1].
- The increasing need for global cooperation, cross-platform governance standards, and ethically grounded constitutions of virtual worlds fills the agenda to allow responsible evolution of the metaverse.

Synthesis of Overall Insight

The findings of this study suggest that while the metaverse presents the transformative potential of changing the very tide of digital interactions, economy, and governance, it simultaneously introduces systemic risks and lacunas. Migrating forthwith calls for a multi-stakeholder approach, including developers, legal experts, policymakers, as well as end-users, to jointly create an inclusive, ethical, and economically feasible metaverse, which is governed by transparent and participative frameworks.

9. Conclusion

Being primarily underlined by blockchain, AI, AR, VR, and decentralized networks, the metaverse opens opportunities that are never seen before in the history of innovation and collaboration. Hence, the term metaverse refers to a new digital paradigm that carries the ability to disrupt industries, economic systems, and social relations. While these opportunities for innovation and collaboration are virtually unprecedented, it also stresses matters that require certain attention. (Zallio & Clarkson, 2023)[1].

In economic terms, the growth of virtual real estate, NFTs, and decentralized finance put into evidence a proof of existence for blockchain-based economies. Yet, issues about fraud-related matters, market volatility, and lack of accessibility have to be dealt with to ensure the participation of all economic agents on equal footing. (Radanliev et al., 2023)[2].

From the social point of view, the metaverse possesses the capability of increasing interaction and hence creativity. On the contrary, serious concerns such as privacy, mental health, and inclusivity threaten to overpower the positive impacts of such virtual environments. Besides that, pertinent ethical issues of data exploitation, users' safety, and monopolization require urgent attention (Garon, 2022)[3].

Arguably, governance is among the most urgent impediments being faced in the metaverse. Decentralized governance structures, such as DAOs, present new avenues for user participation; however, power continues to be concentrated in the hands of the very rich, and a comprehensive legal framework seems inexistent to date to bring forth the realization of equitable and responsible governance. (Garon, 2022)[3].

Stakeholders must come together to establish ethical guidelines and governance structures prioritizing inclusivity, privacy, and user autonomy while embracing the potentials and risks of the metaverse. Research going forward needs to ensure the development of models that secure a balance between technological innovation and responsible stewardship such that the metaverse becomes a positive addition to human society (Zhang, Chen & Liu, 2023)[4].

Recommendations

Based on the findings, the following recommendations are proposed to guide the responsible development of the metaverse:

10. Developing Robust Governance Frameworks

- Governments, corporations, and decentralized entities must collaborate to establish clear legal guidelines addressing privacy, intellectual property, data protection, and user accountability.

- The creation of international standards for decentralized governance models, particularly DAOs, is essential to prevent monopolistic control and ensure democratic decision-making.
- Regulatory bodies should be established to monitor activities within the metaverse and ensure compliance with ethical standards. (Garon, 2022)[3].

11. Promoting Ethical Design Principles

- Developers must prioritize ethical considerations when designing metaverse platforms, ensuring user safety and inclusivity.
- Implementing advanced privacy protection features and ensuring user consent before data collection is crucial.
- Creating mechanisms for monitoring and mitigating harassment, exploitation, and unethical behaviour within virtual environments. (Zhang, Chen & Liu, 2023)[4].

12. Fostering Inclusive Economic Models

- Decentralized financial systems should be designed to promote economic inclusivity and accessibility for marginalized groups.
- Educational platforms within the metaverse should be developed to provide skill-building opportunities for users across different socioeconomic backgrounds. (Radanliev et al., 2023)[2].

13. Enhancing Technological Infrastructure

- Investing in the development of more efficient blockchain systems to reduce energy consumption and enhance transaction speed.
- Promoting the integration of AI-driven technologies to enhance accessibility for users with disabilities. (Zallio & Clarkson, 2023)[1].

14. Encouraging Ongoing Research and Collaboration

- Continued research is needed to understand the psychological impacts of prolonged interaction within virtual environments.
- Collaboration between academia, industry, and government agencies should focus on exploring ethical implications and developing best practices. (Garon, 2022; Zhang, Chen & Liu, 2023)[3,4].

Footnotes (Bluebook Citation Style)

- Matteo Zallio & P. John Clarkson, *Metavethics: Ethical, Integrity and Social Implications of the Metaverse*, 29 Yale J.L. & Tech. 156 (2023).
- Lyubomir Radanliev, De Roure, D., Nurse, J.R.C., & Burnap, P., *The Rise and Fall of Cryptocurrencies: Risk Assessments in Metaverse Ecosystems*, 18 Berkeley Tech. L.J. 204 (2023).
- Jeremy M. Garon, *Legal Implications of a Ubiquitous Metaverse*, 25 Harv. J.L. & Tech. 312 (2022).
- Xia Zhang, Huan Chen & Yu Liu, *Is Metaverse Better than Video Conferencing? Psychological Impact of Immersive Platforms*, 12 J. Virtual Stud. 98 (2023).
- NEAL STEPHENSON, *Snow Crash* (Bantam Books 1992).

AI AND HUMAN RIGHTS IN THE DIGITAL ERA: NAVIGATING THE CROSSROADS OF PROMISE AND PERIL

Siddharth Singh[1]

ABSTRACT:

Artificial Intelligence is no longer a distant promise-it is today's reality, reshaping our world, redrawing the boundaries of industry and society, and all facets of their interaction. As AI's influence surges, it brings with it a double-edged sword: on one side, the shimmering potential to uplift human welfare and fortify rights; on the other, the threat of undermining the very liberties it could enhance. This article examines the complex interplay between AI and human rights, revealing how these technologies can both enable and hinder the path to justice in our digital age. By tracing AI's footprints across vital domains-from privacy and the battle against discrimination to free expression and the pillars of due process- this article offers a comprehensive assessment of AI governance today and suggests pathways towards a world where one complements the other.

The Evolution of AI and Human Rights Discourse

AI has become the "*north-star*" guiding transformation across every sphere of human endeavour, Artificial Intelligence has erupted as a transformative force, weaving itself into the tapestry of daily life-from the hospital ward to the classroom, the courtroom to the digital commons. These systems now shape decisions that impact both, the individual and collective facades of

[1] Student of Law

society. Yet, as these digital sentinels rise, they cast long questions over the human rights frameworks conceived in a pre-digital dawn.

The relationship between AI and human rights is a prism, refracting both hope and hazard. On one facet, AI holds the torch of progress, providing unprecedented opportunities to enhance human capabilities, bridge social inequalities, and advance human rights, developing new avenues for empowerment, equality, and access. As documented in research, "AI-driven assistive technologies enable visually impaired individuals to navigate the world using real-time audio inputs generated by wearable cameras. Similarly, AI applications for individuals with hearing impairments or mobility challenges are revolutionizing access to essential services".[2] Legal AI applications, meanwhile, open doors for marginalized communities, breaking down barriers to justice and opportunity.

But on the flip side, the picture darkens. Without vigilant safeguards, AI can become an invisible oppressor as its algorithms quietly amplify bias, its systems outpacing the laws meant to protect us. As studies warn, AI's unchecked advance risks forging new chains of discrimination, especially for the vulnerable. According to research findings, "AI disproportionally affects the human rights of vulnerable individuals and groups by facilitating discrimination, thus creating a new form of oppression."[3]

This tension-between AI's promise and its risks to human rights has catalyzed global discussions on establishing ethical guidelines and regulatory

[2] 'The Intersection of AI and Human Rights: Ensuring Ethical Standards' (Tata Elxsi) https://www.tataelxsi.com/news-and-events/the-intersection-of-ai-and-human-rights-ensuring-ethical-standards accessed 20 May 2025.

[3] LSE Human Rights, 'Beginning of Artificial Intelligence, End of Human Rights?' (LSE Human Rights Blog, 16 July 2020) https://blogs.lse.ac.uk/humanrights/2020/07/16/beginning-of-artificial-intelligence-end-of-human-rights/ accessed 20 May 2025.

frameworks. International organizations, policymakers, civil societies like those of advocates, and technologists seem to agree on a united, urgent truth: human rights must be the bedrock and foundation of AI governance. As one expert puts it, “Human rights are central to what it means to be human... AI, its systems and its processes have the potential to alter the human experience fundamentally.”[4]

To gain a holistic view of the governance and administration challenges posed by AI, we shall take on a historical overview of its evolution. Digital rights emerged as an extension of fundamental human rights in response to the internet's expansion. According to research, "Digital rights are merely an extension of the rights set out in the Universal Declaration of Human Rights by the United Nations as applied to the online world.”[5] As technology has evolved from basic internet connectivity to sophisticated AI systems, the discourse has similarly progressed from advocating for access to digital technologies toward addressing more complex questions about algorithmic decision-making, data governance, and machine autonomy.

AI and Privacy Rights: The New Battleground

Privacy is one of the elemental concerns, that finds its walls breached by the relentless gaze of AI. AI’s insatiable hunger for personal data-our digital footprints, faces, and even our health- for training and operation, raises profound questions about autonomy and dignity, the scale and nature of data collection, processing, and analysis that is enabled by AI technologies.

The privacy risks posed by AI are multidimensional. According to research, "The right to a private life is threatened by the constant tracking and

[4] Chatham House, 'AI Governance and Human Rights' (2023) https://www.chathamhouse.org/2023/01/ai-governance-and-human-rights accessed 20 May 2025.

[5] ‘What Are Digital Rights?’ (Iberdrola) https://www.iberdrola.com/innovation/what-are-digital-rights accessed 20 May 2025.

surveillance that AI systems use for data collection. The lack of transparency about how AI systems operate creates uncertainty for individuals, whose data can reveal not only their interests but also their vulnerabilities."[6] This core concerns here stems from the significant power imbalance that is created, and where organizations possess extensive knowledge about individuals while the latter remains largely unaware of how their data is being used or for whose benefit.

Biometric identification, especially facial recognition, looms as one such, rather ominous development. Once unimaginable, mass surveillance is now a reality. In the European Union AI Act, "real-time biometric identification systems in public spaces are generally prohibited, except in cases deemed necessary for national security or criminal investigations."[7] However, these exceptions "leave room for potential overreach, raising the risk of government-led surveillance programs that may disproportionately impact marginalized groups."[8]

The collection and processing of healthcare data, too, finds itself at the "*crosshairs*". AI processes our most sensitive data, demanding robust protections. Research emphasizes the need for careful assessment of AI's impact in healthcare contexts-especially for the elderly and vulnerable: "AI solutions may alleviate the health workforce crisis and contribute to the highest attainable standard of physical and mental health for all in an ageing

[6] European Network of National Human Rights Institutions, 'Key Human Rights Challenges' (ENNHRI 2022) https://ennhri.org/ai-resource/key-human-rights-challenges/ accessed 20 May 2025.

[7] Emilio Dávila, 'EU AI Act and Human Rights Compliance' (EmilDAI, 15 March 2023) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025.

[8] Emilio Dávila, 'EU AI Act and Human Rights Compliance' (EmilDAI, 15 March 2023) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025.

Europe. At the same time, however, these solutions may create new and unforeseen human rights challenges."[9]

Even International human rights bodies have rung an urgent call for action to address AI's privacy implications. The UN High Commissioner for Human Rights has called for immediate action, stating: "Artificial intelligence can be a force for good, helping societies overcome some of the great challenges of our times. But AI technologies can have negative, even catastrophic, effects if they are used without sufficient regard to how they affect people's human rights."[10] The High Commissioner specifically emphasized the need for a moratorium on AI systems that pose serious risks to human rights until adequate safeguards are established.

The notorious "black box" problem-AI's opacity-casts a long shadow over privacy. Research highlights the importance of transparency principles: "Transparency and explainability allow individuals affected by AI to be informed in a timely, comprehensive and clear manner about issues concerning the use of their personal information in AI processes and the possible consequences of the specific reasons behind such use."[11] Without transparency and explainability, individuals are left powerless to understand or challenge decisions made about them, eroding trust and accountability.

Discrimination and Bias: The Echoes of Inequality

[9] Philip Czech, Lisa Heschl, Karin Lukas, Manfred Nowak, and Gerd Oberleitner, 'A Human Rights-Based Approach to Artificial Intelligence in Healthcare: A Proposal for a Patients' Rights Impact Assessment Tool' (SSRN Paper 4861569, 3 June 2024) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4861569 accessed 20 May 2025.

[10] Office of the UN High Commissioner for Human Rights, 'Artificial Intelligence Risks to Privacy Demand Urgent Action – Bachelet' (OHCHR 2021) https://www.ohchr.org/en/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet accessed 20 May 2025.

[11] Office of the UN High Commissioner for Human Rights, 'Privacy Key in Processing Personal Data in AI' (OHCHR 2023) https://www.ohchr.org/en/press-releases/2023/10/privacy-key-processing-personal-data-ai-un-expert accessed 20 May 2025.

One of the most pressing human rights issues related to AI is its potential to mirror-and-magnify existing forms of discrimination. Because AI systems are frequently trained on historical data, they tend to mirror—and sometimes worsen—societal biases, resulting in discriminatory impacts that fall most heavily on marginalized communities. This situation poses significant challenges to the principles of equality and non-discrimination that are fundamental to international human rights standards.

Discriminatory outcomes in AI systems can manifest horizontally, across various domains, including hiring processes, law enforcement, financial services, and healthcare. A recent instance highlighting this issue: "Plaintiff alleges Workday's algorithmic decision-making tools discriminate against job applicants who are African-American, over the age of 40, and/or are disabled."[12] In this case, the court notably refused to "draw an artificial distinction between software decision-makers and human decision-makers," recognizing that automated systems should not be exempt from anti-discrimination laws."[13]

Algorithmic bias has various points-of-genesis. Research identifies that "One fundamental challenge is that bias in AI is often deeply embedded in historical data. Even if an AI system is designed to be 'neutral,' it may still replicate past patterns of discrimination due to the historical inequalities reflected in training datasets."[14] This creates a vicious feedback loop where

[12] 'Artificial Discrimination: AI Vendors May Be Liable for Hiring Bias' (Clark Hill, 10 June 2023) https://www.clarkhill.com/news-events/news/artificial-discrimination-ai-vendors-may-be-liable-for-hiring-bias-in-their-tools/ accessed 20 May 2025.

[13] 'Artificial Discrimination: AI Vendors May Be Liable for Hiring Bias' (Clark Hill, 10 June 2023) https://www.clarkhill.com/news-events/news/artificial-discrimination-ai-vendors-may-be-liable-for-hiring-bias-in-their-tools/ accessed 20 May 2025.

[14] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025

AI systems perpetuate historical injustices under the veneer of objective, data-driven decision-making.

Facial recognition technologies, for one, have proven particularly problematic regarding racial and ethnic bias. Studies have shown that "Technologies such as facial recognition and language modelling have shown prejudice against racial and ethnic minorities, leading to injustices such as false arrests and accusations."[15] These technological shortcomings can have severe consequences for individuals subjected to biased systems, particularly in high-stakes contexts like criminal justice.

Predictive policing is another cautionary tale; another domain where AI bias raises serious human rights concerns. While some studies describe apparent successes in crime reduction through predictive policing programs[16], these systems often rely on historically biased crime data that reflects patterns of over-policing in minority communities. Without careful extermination of these underlying data biases, predictive policing risks reinforcement of discriminatory practices.

These concerns are furthered by the opacity of many AI systems, making it difficult to detect and address bias. Research observes that "while the AI Act mandates transparency and explainability, it does not fully address the issue of algorithmic accountability. If an individual is denied a job or a loan due to an AI-driven decision, what legal recourse do they have?"[17] This lack of

[15] ENNHRI, 'Key human rights challenges of AI' (ENNHRI,) https://ennhri.org/ai-resource/key-human-rights-challenges/ accessed 20 May 2025

[16] Ibrahim Raji and Damilola Bartholomew Sholademi, 'Predictive Policing: The Role of AI in Crime Prevention' (2024) 13(10) International Journal of Computer Applications Technology and Research 66 https://ijcat.com/archieve/volume13/issue10/ijcatr13101006.pdf accessed 20 May 2025

[17] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025

accountability creates significant barriers to justice for victims of algorithmic discrimination.

Freedom of Expression and Information in the Age of AI

The proliferation of AI technologies has profound implications that extends to the very heart of democracy: freedom of expression and information. Algorithms now dictate, curate, moderate, and sometimes silence what we see, hear, and say online, wielding unprecedented power over the flow of information.

Content moderation, is one of the tasks that has seen proliferate automation. Major online platforms are rapidly adopting AI algorithms to detect and remove content that violates platform policies or legal standards. While this automation allows for processing massive volumes of content, it raises concerns about algorithmic censorship. According to research, "AI plays a growing role in content moderation on digital platforms... While AI can enhance and improve efficiency, it also poses risks to due process, legal fairness, and the suppression of lawful speech."[18]

At the heart of this issue, lies the AI's limited, or at times absence, of an ability to grasp or detect nuance, context, and cultural specificity-essential elements for evaluating expression. As research notes, "Many social media platforms rely on AI algorithms to detect and remove harmful content, such as hate speech and misinformation. While combating online abuse is essential, AI-driven censorship mechanisms often lack significance, leading to the unfair suppression of legitimate political speech, activism, and

[18] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025

disagreement.”[19] This creates a risk of over-censorship where protected speech is incorrectly flagged and removed.

AI's role in information dissemination and curation also raises concerns about manipulation of public discourse. Recommendation algorithms possess the risk of trapping us in echo chambers "filter bubbles, narrowing our perspectives and subtly shaping our choices, and at large, public discourse, through the determination of what news, entertainment, or social media the content users see. Research observes that "systems that have algorithms with addictive designs or that create echo chambers, such as some social media platforms, can influence the ability to freely make choices and decisions without coercion or manipul[ation].”[20]

AI-generated content presents another frontier for freedom of expression concerns. Advanced language models can now generate highly convincing text that mimics human writing, raising questions about authenticity, attribution, and accountability in public discourse. Research highlights how "the protection of freedom of expression has been deeply impacted by the development of digital technologies and, in particular, by the spread of artificial intelligence systems.”[21] With advancement, the line between human and AI-generated content becomes increasingly blurry, potentially undermining trust in information and creating new vectors for disinformation campaigns.

[19] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025

[20] ENNHRI, 'Key human rights challenges of AI' (ENNHRI,) https://ennhri.org/ai-resource/key-human-rights-challenges/ accessed 20 May 2025

[21] Giovanni De Gregorio, ‘Artificial Intelligence and Freedom of Expression’ (SSRN, 9 June 2023) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4736744 accessed 21 May 2025.

Legal frameworks protecting freedom of expression are encountering major obstacles in the context of artificial intelligence. Traditionally, these frameworks have concentrated on limiting government censorship. However, the rise of AI-driven platforms means that private companies, through their algorithmic systems, now effectively function as gatekeepers of online expression. Research notes that "governance of these expressions is increasingly mediated by artificial intelligence systems implemented by states and business actors such as in the case of content moderation."[22]

AI in Judicial Systems and Law Enforcement: Justice at a Crossroads

The deployment of AI technologies in justice systems and law enforcement settings raises profound concerns regarding due process rights, fair trial guarantees, and the rule of law. Although these applications offer the potential for increased efficiency and enhanced objectivity, they also pose new threats to the fundamental rights that underpin justice systems.

Algorithmic decision-making in judicial contexts has been found to potentially undermine established due process protections: "Regarding Judge Henry J. Friendly's procedural due process principles of the U.S. Constitution, decisions produced using AI appear to violate all but one or two of them. For instance, AI systems may provide the right to present evidence and notice of the proposed action, but do not provide any opportunity for meaningful cross-examination, knowledge of opposing evidence, or the true reasoning behind a decision."[23]

[22] Giovanni De Gregorio, 'Artificial Intelligence and Freedom of Expression' (SSRN, 9 June 2023) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4736744 accessed 21 May 2025.

[23] Chris Chambers Goodman, 'AI, Can You Hear Me? Promoting Procedural Due Process in Government Use of Artificial Intelligence Technologies' (2024) 28(4) Rich J L & Tech 700 https://scholarship.richmond.edu/jolt/vol28/iss4/3 accessed 20 May 2025.

Predictive justice tools, which attempt to assess recidivism risk or inform sentencing decisions, have prompted particular concern. According to research, "AI-powered predictive policing tools, for example, are often based on historical crime data that may reflect racial and socioeconomic biases. If courts increasingly rely on AI-generated risk assessments for sentencing decisions, there is a danger that individual rights to a fair trial may be undermined."[24] These tools risk replacing individualized judicial assessment with statistical generalizations that may penalize individuals based on demographic characteristics rather than personal conduct.

Law enforcement through AI, especially in predictive policing raises similar concerns. While some jurisdictions report that predictive policing strategies reduced crime rates[25], these systems rely on historical crime data that often reflects discriminatory policing practices. The risk is creating a feedback loop that reinforces existing patterns of over-policing in marginalized communities. Predictive policing enables "law enforcement agencies to allocate resources more effectively by identifying hotspots of criminal activity"[26], but this data-driven approach may concentrate enforcement in communities that are already disproportionately targeted.

Facial recognition technology in law enforcement presents particularly acute human rights risks, demonstrating significant accuracy disparities across demographic groups, especially with higher error rates for the marginalized like women and people with darker skin tones. When deployed in public

[24] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025

[25] Ibrahim Raji and Damilola Bartholomew Sholademi, 'Predictive Policing: The Role of AI in Crime Prevention' (2024) 13(10) International Journal of Computer Applications Technology and Research 66 https://ijcat.com/archieve/volume13/issue10/ijcatr13101006.pdf accessed 21 May 2025.

[26] Ibrahim Raji and Damilola Bartholomew Sholademi, 'Predictive Policing: The Role of AI in Crime Prevention' (2024) 13(10) International Journal of Computer Applications Technology and Research 66 https://ijcat.com/archieve/volume13/issue10/ijcatr13101006.pdf accessed 21 May 2025.

spaces for surveillance purposes, facial recognition technology effectively eliminates anonymity and creates chilling effects on protected activities like political protest. The UN High Commissioner for Human Rights has called for "a moratorium on the sale and use of artificial intelligence (AI) systems that pose a serious risk to human rights until adequate safeguards are put in place."[27]

Regulatory Frameworks and Ethical Guidelines for AI Governance

As the *AI-tide rises* against the *banks of society,* governments, international bodies, and private actors have scrambled to erect dams-regulatory frameworks and ethical codes-to channel its force while mitigating its risks to human rights and democratic values.

The European Union's AI Act is the pioneering legislation globally, that aims to regulate and govern AI utilization. According to research, the Act "represents a significant step forward in regulating AI technologies in a manner that respects human rights and democratic values."[28] The framework adopts a risk-based approach, imposing stricter requirements on high-risk AI applications that could significantly impact fundamental rights. The Act explicitly aligns with the European Convention on Human Rights (ECHR), reflecting the EU's commitment to rights-based AI governance.

International organizations have also developed influential ethical frameworks. The OECD AI Principles outline five key pillars: "inclusive growth, sustainable development, and human well-being; respect for the rule

[27] Office of the UN High Commissioner for Human Rights, 'Artificial Intelligence Risks to Privacy Demand Urgent Action – Bachelet' (OHCHR 2021) https://www.ohchr.org/en/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet accessed 20 May 2025.

[28] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025.

of law, human rights, and democratic values (including fairness and privacy); transparency and explainability; robustness, security, and safety; and accountability."[29] Their significant traction globally, has acted as 'guardrails', informing national AI strategies and corporate policies.

The United Nations has time-and-again, underscored the importance of human rights as a foundation for AI governance. The UN High Commissioner for Human Rights has called for "adequate safeguards" before deploying AI systems that pose serious risks to human rights.[30] The statement emphasizes that "the complexity of the data environment, algorithms and models underlying the development and operation of AI systems, as well as intentional secrecy of government and private actors are factors undermining meaningful ways for the public to understand the effects of AI systems on human rights and society."[31]Regional human rights mechanisms have also addressed AI governance. The Council of Europe recognizes both AI's potential to "significantly enhance the protection and promotion of human rights" and the "serious risks" it poses, including "discrimination, gender inequality, threats to democratic processes, infringements on human dignity and autonomy, and the misuse of AI by States for repressive purposes."[32] The organization is developing a Handbook on Human Rights and Artificial Intelligence to provide practical guidance on upholding human rights in the AI era.

[29]

[30] Office of the UN High Commissioner for Human Rights, 'Artificial Intelligence Risks to Privacy Demand Urgent Action – Bachelet' (OHCHR 2021) https://www.ohchr.org/en/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet accessed 20 May 2025.

[31] Office of the UN High Commissioner for Human Rights, 'Artificial Intelligence Risks to Privacy Demand Urgent Action – Bachelet' (OHCHR 2021) https://www.ohchr.org/en/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet accessed 20 May 2025.

[32] Council of Europe, 'Human Rights and Artificial Intelligence (CDDH-IA)' (Council of Europe, undated) https://www.coe.int/en/web/human-rights-intergovernmental-cooperation/intelligence-artificielle accessed 21 May 2025.

Despite this proliferation of frameworks, significant lacunae persist. Research notes that "many sets of AI governance principles produced by companies, governments, civil society and international organizations do not mention human rights at all. This is an error that requires urgent correction."[33] This highlights the need to explicitly ground AI governance in established human rights principles rather than creating parallel ethical frameworks disconnected from legal human rights obligations.

Implementation and enforcement, particularly, prevail as challenges in AI governance. Research observes that while the EU AI Act establishes important safeguards, "its practical enforcement will determine whether it truly safeguards fundamental rights or simply sets aspirational guidelines."[34] Effective governance requires not only well-designed rules but also robust oversight mechanisms, including independent regulatory bodies with sufficient technical expertise and enforcement authority.

Balancing Innovation with Human Rights Protection

A central point of tension in AI governance lies in striking the right balance between promotion of beneficial innovation and safeguarding strong human rights protections. This challenge mirrors broader discussions about how societies can leverage technological progress while mitigating its potential risks.

The case for innovation is compelling. AI technologies offer significant opportunities to advance human rights and well-being across various domains. Research highlights how AI "empowers organisations to unlock new opportunities, drive innovation, and deliver enhanced customer

[33] Chatham House, 'AI Governance and Human Rights' (2023) https://www.chathamhouse.org/2023/01/ai-governance-and-human-rights accessed 20 May 2025.

[34] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025.

experiences," including through "enhanced data analytics and insights," "automation of repetitive tasks," and "personalised customer experiences."[35] Similarly, AI "has the potential to advance human rights by fostering inclusivity, accessibility, and justice," particularly for marginalized groups.[36] However, innovation without adequate guardrails poses substantial risks. UN High Commissioner for Human Rights Michelle Bachelet warned: "We cannot afford to continue playing catch-up regarding AI-allowing its use with limited or no boundaries or oversight, and dealing with the almost inevitable human rights consequences after the fact."[37] This reactive approach to governance has repeatedly proven insufficient as technologies advance faster than regulatory frameworks can adapt.

Some stakeholders frame human rights protections as potential barriers to innovation, creating a false dichotomy between technological progress and rights safeguards. However, research suggests a more nuanced perspective: "While AI promises transformative benefits, its misuse can exacerbate societal disparities. Striking a balance between innovation and ethical oversight is key to ensuring AI serves as a tool for empowerment rather than oppression."[38] This framing recognizes that sustainable innovation requires public trust, which in turn depends on responsible development and deployment practices.

[35] 'How Does AI Impact Digital Transformation?' (YourShortlist) https://yourshortlist.com/how-does-ai-impact-digital-transformation/ accessed 20 May 2025.
[36] 'The Intersection of AI and Human Rights: Ensuring Ethical Standards' (Tata Elxsi) https://www.tataelxsi.com/news-and-events/the-intersection-of-ai-and-human-rights-ensuring-ethical-standards accessed 20 May 2025.
[37] Office of the UN High Commissioner for Human Rights, 'Artificial Intelligence Risks to Privacy Demand Urgent Action – Bachelet' (OHCHR 2021) https://www.ohchr.org/en/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet accessed 20 May 2025.
[38] 'The Intersection of AI and Human Rights: Ensuring Ethical Standards' (Tata Elxsi) https://www.tataelxsi.com/news-and-events/the-intersection-of-ai-and-human-rights-ensuring-ethical-standards accessed 20 May 2025.

Proactive approaches incorporating human rights considerations from the earliest stages of design and development can help reconcile innovation with rights protection. This "human rights by design" approach integrates rights considerations into the technical development process rather than treating them as external constraints. UNESCO's "human rights approach to AI" includes principles such as "proportionality and do no harm," "safety and security," and "right to privacy and data protection."[39]

Multi-stakeholder governance models represent a dynamic and all-round approach to balancing innovation and rights protection. The importance of "multi-stakeholder" approaches to AI governance[40] recognizes that effective frameworks require input from diverse perspectives, including technical experts, rights advocates, affected communities, and policymakers. This collaborative approach can help identify potential rights impacts early in the development process while preserving space for beneficial innovation.

Conclusion: Toward a Human Rights-Based Approach to AI Governance

As AI continues its venture and advance into every domain of human life, a human rights-based approach to their governance emerges as both a normative imperative and a practical necessity. This approach recognizes that established human rights principles-though developed before the digital age-provide essential guidance for ensuring AI systems respect human dignity and fundamental freedoms.

This article evinces throughout. how AI systems can both enhance and potentially undermine human rights. As the UN High Commissioner for

[39] UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' (UNESCO, 2021) https://www.unesco.org/en/artificial-intelligence/recommendation-ethics accessed 20 May 2025.

[40] UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' (UNESCO, 2021) https://www.unesco.org/en/artificial-intelligence/recommendation-ethics accessed 20 May 2025.

Human Rights observed: "Artificial intelligence can be a force for good, helping societies overcome some of the great challenges of our times. But AI technologies can have negative, even catastrophic, effects if they are used without sufficient regard to how they affect people's human rights."[41] This implies an urgent need for governance frameworks that maximize AI's positive contributions while establishing effective safeguards against its risks.
A human rights-based approach to AI governance offers several advantages over alternative and entirely technical frameworks. First, it builds on established legal principles with widespread international consensus, providing a shared vocabulary and normative foundation across diverse cultural and political contexts. Second, it recognizes the indivisibility and interdependence of rights, acknowledging that AI systems may simultaneously impact multiple rights such as privacy, non-discrimination, and freedom of expression. Third, it emphasizes the importance of accountability mechanisms and access to remedy when rights violations occur.

Implementing a human rights-based approach to AI governance requires action across multiple levels. At the international level, research recommends that "human rights [should be] the foundation for AI governance in future."[42] This includes developing international standards and cooperative mechanisms to address transboundary challenges posed by global AI systems.

At the national level, governments must ensure that domestic legal frameworks adequately protect human rights in the context of AI

[41] Office of the UN High Commissioner for Human Rights, 'Artificial Intelligence Risks to Privacy Demand Urgent Action – Bachelet' (OHCHR 2021) https://www.ohchr.org/en/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet accessed 20 May 2025.
[42] Chatham House, 'AI Governance and Human Rights' (2023) https://www.chathamhouse.org/2023/01/ai-governance-and-human-rights accessed 20 May 2025.

development and deployment. This includes not only adopting AI-specific regulations where necessary but also ensuring that existing human rights protections extend effectively to AI contexts. Research emphasizes the importance of "independent oversight mechanisms, stronger legal remedies, and clear enforcement strategies" to give practical effect to rights protections.[43]

Private sector players, particularly those developing and deploying AI systems, bear significant responsibility for respecting human rights. Courts have refused to "draw an artificial distinction between software decision-makers and human decision-makers," affirming that "delegating a function to an automated or AI system does not insulate the employer or vendor from liability for the decisions made by those tools."[44] This principle of accountability should inform corporate governance of AI systems.

Looking ahead, several key priorities stand out for advancing a human rights-based approach to AI governance:

1. Strengthening transparency requirements to address the "black box" issue, which undermines meaningful accountability in AI systems.
2. Developing more effective remedies for individuals who are harmed by AI systems that violate their rights.
3. Enhancing participatory governance mechanisms to ensure that affected communities have a meaningful role in shaping AI policy.
4. Building both technical and governance capacity, especially in regions with fewer resources, to support effective oversight of AI systems.

[43] Dr Srabonty Das Gupta, 'The EU AI Act and Its Adherence to the European Convention on Human Rights' (EMILDAI, 17 May 2024) https://emildai.eu/the-eu-ai-act-and-its-adherence-to-the-european-convention-on-human-rights/ accessed 20 May 2025.

[44] 'Artificial Discrimination: AI Vendors May Be Liable for Hiring Bias in Their Tools' (Clark Hill, 10 June 2023) https://www.clarkhill.com/news-events/news/artificial-discrimination-ai-vendors-may-be-liable-for-hiring-bias-in-their-tools/ accessed 20 May 2025.

5. Creating technical standards and certification processes that respect human rights, in order to integrate these considerations into mainstream AI development.
6.

As researchers note, AI is "transforming our society, profoundly affecting how we live, work, and interact.”[45] Ensuring this transformation advances, rather than undermining, human dignity and fundamental rights requires concerted effort across stakeholders and governance levels. A human rights-based approach provides the principled foundation and practical guidance needed to navigate this transformative period in human history.

[45] Council of Europe, ‘Human Rights and Artificial Intelligence (CDDH-IA)’ (Council of Europe, undated) https://www.coe.int/en/web/human-rights-intergovernmental-cooperation/intelligence-artificielle accessed 21 May 2025.

ALGORITHMIC JUSTICE: THE ROLE OF AI IN MODERN COURTROOMS

Sheetal Kumari And Prerna Singh[1]

Abstract

Is it possible that AI can take over human judges, or can we expect AI software's to play the classic traditional role of judges in the near future?
Well, that is what we are going to discuss in this research paper to better understand this, we need to 1st look at the judicial process and how it operates and how AI will function in this system. In the recent years, the use of artificial intelligence in legal system has expanded because of its efficiency to use and is also impartial in comparison to human judges. But it also has certain limitations because it is based on big data and also works on algorithms rather than morality, which is essential aspect to consider while providing justice and also it is based on computing power rather than organic intelligence. In the current scenario, we should keep aside the possibility of advances of radical technology and we should acknowledge that the role of AI is to support the human judges, not replace them.

The main aim of this research paper is to study how AI impacts the legal system, and how it effects in context of algorithmic justice. And also if the software is developed in such a manner to replace the functions of human judges, there will still remain and uncertainty of whether these algorithms will work in relation to Indian society. Henceforth, AI may be more capable and easy to use, but it will still remain alien in fundamental ways. As of 2025, there has been a significant backlog of cases in the Indian judiciary across

[1] students of Ba.Llb Iind Year (Ivth Semester)

various levels of the court. In Supreme Court approximately 80,221 cases were pending as of January 23, 2024, over 5,800,000 cases were collectively pending by January 2025 in High-court and about 45 million cases were pending in district and subordinate courts as of November 2024. These numbers can be reduced with the effective use of AI by focusing on three main areas. Firstly, legal research to analyse the reasons behind the pending cases. Secondly, legal reasoning given by the judges and lastly, the predictable outcome from the precedents. Overall, this paper analyses the role of AI in modern courtrooms while evaluating the feasibility of AI in current Indian judicial system. Since this paper discusses the role of artificial intelligence in shaping society and law it is valuable not only for the legal professionals and Lord Students, but also for engineering students and engineers who are seeking to advance AI technology in the legal sector as well as for those who will be the consumers of this technology.

Keywords: Artificial Intelligence (AI), Judicial Process, Legal System, Limitations, Support vs. Replacement, Indian Judiciary, Pending Cases, Legal Research, Precedents, AI in Legal Sector.

1. Introduction

Nowadays, machines have become an alternative to human labor or workforce. For instance, if we want to visit a place where we aren’t sure of the directions, Google Maps are of great use, and they also indicate the facilities that are on the way through the algorithm embedded in it. We all are dependent on AI in different ways, which makes our lives easier. In recent headlines, the robot “Justice” has been in debate. There are claims that algorithms fed into AI can predict outcomes accurately and that we won’t need human judges anymore.

In Indian judicial settings, there are a huge number of pending cases that are unresolved due to the lack of judges in proportion to the number of cases. AI has the potential to assist with this issue by systematically arranging backlog cases. They are updated online via a website—NJDG (National Judicial Data Grid)[2]—making it convenient for courtroom proceedings and easily accessible by laypersons.

According to article 6 of ECHR[3] and the ethical guidelines. There should be a transparent procedure where parties to the proceedings are treated equally and by a well-founded judgment. Section 3 particularly talks about reducing judicial complexity, which should be substantiated by providing a level playing field to the litigants. If human judges are replaced by AI, then there are possibilities that data automation also contains legally incorrect decisions or biases, which will ultimately reduce the quality of AI judgments. Thereby it will violate the provisions of ECHR.

To incorporate AI in the judicial system fully, it should be trained to explain its outcomes and should not be merely based on precedents.

Currently, AI can be of great use in advising, translating, legal interpretation, and organizing overall information. Moreover, the facts and circumstances vary from case to case, which can be solved by rationality, not merely on the basis of precedents or provisions.

I. **Hypothesis**

a) Primary Hypothesis

Although artificial intelligence has the potential to assist judges in decision-making and enhance procedural efficiency, it cannot fully replace the role of

[2] NJDG (National Judicial Data Grid)
[3] ECHR (European Convention on Human Rights)

human judges due to ethical, legal, and constitutional concerns—particularly those concerning transparency, accountability, and fairness of judicial proceedings.

b) Sub Hypothesis

Suppose the use of AI in courtrooms is not effectively regulated. In that case, it may lead to violations of fundamental rights and infringing upon constitutionally guaranteed fundamental rights, thereby posing a threat to individual liberties.

II. Evolution of AI in the legal system

a) Global development -

□USA

The USA first used AI in a legal context in early 2010 for legal research and case prediction. AI platforms like Ross intelligence were developed to assist lawyers while researching case laws and legal precedent. AI technologies have also been used to automate the process of reviewing electronic documents related to litigation, which ultimately reduced the time and cost associated with manual work.

□United Kingdom

AI began in the United Kingdom in mid-2010 to review a large volume of contracts and enhance accuracy. For instance, a top-20 global law firm used AI for contract review in May 2022.

□China

In late 2010, AI was integrated into China's legal system as a scheme of the smart courts initiative, where China implemented AI in digital judicial proceedings. AI technologies have also been introduced to handle routine cases like e-commerce disputes, reduce human workload, assess the authenticity of evidence presented during trials, and streamline case proceedings.

□Estonia-

Introduced in 2019, Estonia announced initiatives to develop AI-driven judges to arbitrate small-claim disputes involving amounts that are less than €7000 to reduce the burden on human judges and advance case resolutions.

□European Union

In 2021, the EU established guidelines to regulate the use of AI in the legal sector. The Artificial Intelligence Act of 2021 was introduced to balance innovation with the protection of fundamental rights.

b) Indian Context-

AI was incorporated into the Indian legal system in the early 2020s by introducing SUPACE (Supreme Court portal for assistance in court efficiency)[4] to assist the judges by providing AI-automated research and summarization tools to enhance decision-making. Another significant development is SUVAAS, an AI-based language translation tool developed to make legal documents and judgments easily accessible in different regional languages.

2. **Artificial intelligence effects on the modern courtroom Procedures**

Due to persistent delays and challenges in ensuring speedy trials, many countries view the judiciary as vulnerable. These uncertainties have led to debates about the effectiveness of the traditional legal system. As a result, artificial intelligence is increasingly being used to stimulate certain judicial functions. While the human judge and jury, AI is being deploy to assist in improving accuracy, efficiency and access to till now, it is not intended to

[4] SUPACE (Supreme Court Portal for assistance in court efficiency)

replace the human judges, but only to supplement the administrative burden and streamlining tasks.

AI can be beneficial in a lot of ways as it can analyze various angles of legal activities. It is impartial, precise, and transparent. But each legal case is unique and needs to be decided according to rationality and experience, which is possible only in the case of human judges and attorneys. Due to the input of data of precedent or the provisions, AI can give judgment accordingly, but it will lack the necessary emotional and resonance components, which is the essence of judicial justice.

Globally, AI is being used in multiple ways. For instance, in the United States, COMPAS[5] has been used to assist judges in bail and sentencing cases. These tools are conflicting and controversial due to concerns of social and racial bias.

In China, AI is applied in smart courtrooms where systems handle civil and commercial disputes, especially those arising from the e-commerce sector. They also help in case management and generate initial drafts of management. Even after all these technological advancements, the role of AI in modern courtrooms is to provide an aid rather than a replacement. It is important to recognize that AI is very useful and economical, but it cannot address all the legal system's challenges.

In India, AI has been used in several key sectors, which include legal research and document analysis, where it assists lawyers and judges in finding relevant case laws; case management systems, which help courts to track and manage the pending cases; and lastly, language translation tools and AI-assisted research tools. It has also been used to digitalize court records and enhance

[5] COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)

public access. One such example is of digitization projects undertaken by the Delhi High Court, which uses AI to process the documents, making large volumes of judicial data easily retrievable and searchable.

3. Advantages of Incorporating AI in Courtrooms

The popularity of AI has increased with the advent of ChatGPT, a chart bot developed by Open AI. By December 2022 and January 2023, it had accumulated over 100 million users. Its application was first used in February 2023 by a Colombian judge. Several questions were raised about the AI chatbot and its relevance in judicial proceedings. In India, it was first used by the Punjab and Haryana High Court, where a bench led by Justice Anup Chitkara took the help from an assistant of a chatbot while giving the verdict of bail application of the accused. It can be clearly stated that the usage of AI is gradually gaining acceptance worldwide, even by judicial institutions.

A. Efficiency and speed

If AI is employed in the courtroom, it will increase efficiency in the legal proceedings, reducing labor-intensive administrative tasks review, case management, legal research, etc. Automated systems can process big data and extract relevant information quickly.

For example, in Singapore, AI has been used to sort legal documents and streamline case management and document analysis, thus reducing the time of lawyers, judges, and clerks.

B. Cost reduction and Access to justice

Many individuals cannot afford the high cost of legal services, which is also a persistent barrier to justice. AI has the potential to reduce the operational costs in both the private and public legal sectors. AI, when trained to do document drafting, can cut down the need for human labor, thereby lowering costs for both clients and legal institutions.

For example, DoNotPay, an AI-driven legal service platform, generates legal documents and offers automated advice, making legal processes more accessible to marginalized groups.

C. **Impartial**

Human bias can often influence judicial outcomes, especially in cases involving gender, class, economic status, race, etc. AI systems, when trained properly, can mitigate such biases by relying on data-driven decisions rather than subjective human judgment.

For example, in the US, the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) system uses machine learning algorithms to assess various factors, such as the likelihood of a defendant re-offending, and on the basis of that, it generates a risk score. However, it has been criticized for perpetuating racial biases in its application.

D. Real-time Translation

AI real-time translation can be beneficial in jurisdictions involving multiple languages or diverse populations, ensuring that all parties have equal access to justice in legal proceedings.

For example, in Europe, AI translation tools are used in quotes to provide real-time translation of witness testimony and legal documents. It helps non-native speakers and reduces language barriers by providing access to justice.

4. **Disadvantages and challenges of AI in Judiciary**

Integrating artificial intelligence into law firms worldwide could significantly alter the nature of legal work. The implementation of artificial intelligence (AI) in the legal sector brings a range of challenges that require careful consideration and management. Key aspects such as compliance with legal regulation (e.g., the AI Act), data protection (e.g., GDPR[6]), ethical concerns, and accountability. A recent example includes the Punjab & Haryana High

[6] GDPR (General Data Protection Regulation)

Court utilizing the AI tool ChatGPT while considering a bail application. Experts in the field believe this marks only the beginning, anticipating a broader role for AI in accelerating case resolutions and improving the efficiency of the justice delivery process.

Nevertheless, critical questions remain: How far can AI be involved in the judicial system? Can it truly take the place of human judgment in decisions that affect people's rights and freedoms?

A. Algorithmic bias: AI models may reflect biases from historical case data.

As artificial intelligence becomes more integrated into various sectors, including the legal system, concerns about algorithmic bias and potential discrimination have intensified. Algorithmic bias occurs when AI consistently produces unfair outcomes that disproportionately affect people from marginalized or less represented communities.

Ethical and Human Rights Concerns Linked to Biased AI

1. Violation of the Right to Non-Discrimination:

- When AI discriminates against individuals based on race, gender, or socioeconomic status, it undermines the fundamental human right to equal treatment (European Commission, 2020).
- Such bias can strengthen existing prejudices within society and result in unfair legal outcomes (Crawford & Schultz, 2013).

2. Invasion of Privacy:

- AI systems influenced by bias may make judgments using sensitive personal information, such as racial background or gender identity (European Commission, 2020).
- This raises serious privacy concerns and highlights the need for strict safeguards around personal data.

3. Issues of Transparency and Accountability:

- The opaque nature of AI decision-making can make it difficult to hold systems or their developers accountable (Crawford et al., 2016).
- It is essential for individuals to understand how and why legal decisions are made, particularly when those decisions have major consequences on their lives.
- Vulnerable and marginalized communities often bear the brunt of biased AI decisions, exacerbating existing disparities (Law, 2018).
- This highlights the need to safeguard the rights of these communities through ethical AI practices
-

B. Lack of transparency: AI decisions may not be explainable or challengeable.

One of the most pressing challenges in using artificial intelligence within the judicial system is the lack of transparency in how these systems arrive at their conclusions. Many AI models—particularly those based on complex machine learning algorithms—function as "black boxes," meaning their internal decision-making processes are not easily understandable, even by experts. This opacity poses a serious concern when such systems are used in matters that directly impact human rights, such as granting bail, sentencing, or case prioritization.

AI systems often rely on large datasets and intricate statistical patterns, which can make their decisions difficult to interpret in human terms. For example, an AI tool may assess the risk of an accused person fleeing bail based on past data, but it might not be clear which factors were most influential in its assessment. This lack of clarity makes it difficult for lawyers, judges, or affected individuals to understand the rationale behind a decision.

A fair judicial process requires that all parties understand the case against them and have an opportunity to respond. AI-driven decisions, if not explainable, threaten this foundational principle. Without transparency, there is no guarantee that AI is operating without bias, making legally sound decisions, or even complying with constitutional values.

C. Accountability Issues: which is responsible when AI makes errors in legal decisions?

AI systems lack the fundamental human qualities of empathy and emotional intelligence that are essential for understanding the human dimensions of legal disputes. Justice Gavai explicitly stated that "the essence of justice often involves ethical considerations, empathy, and contextual understanding elements that remain beyond the reach of algorithms". Challenging legal cases often demand moral reasoning and value-based judgments that go beyond what can be captured through calculations or mathematical formulas. Human judges bring life experience, cultural understanding, and societal context to their decisions that AI cannot replicate regardless of its programming sophistication.

D. Privacy & Data security: Al relies on large datasets, raising concerns about data misuse.

As organizations increasingly depend on AI to handle sensitive information, protecting privacy and ensuring data security have become critical priorities. Unlike traditional systems, AI presents distinct challenges because it can autonomously process, analyse, and learn from data that can sometimes lead to unexpected outcomes—for example, an AI model might unintentionally combine information from multiple sources in a way that reveals the identity of individuals or discloses personal details.

Additionally, securing data within AI systems is vital, as these models can be vulnerable to various cyber threats. Attackers may manipulate the AI’s behaviour or extract confidential information from its training data. Unauthorized access to either the AI algorithms or the data they use can lead to severe breaches, harming individuals' privacy and damaging an organization's reputation and compliance status.

Laws such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) require organizations to adopt strong privacy and security measures. These regulations emphasize that protecting personal data isn't optional—it's a legal and operational necessity in today's digital environment.

E. Legal & Ethical concerns: AI replacing human judgment could threaten due process.

The increasing use of AI in judicial and legal decision-making raises significant legal and ethical questions.

Threat monitoring and detection

Justice systems are fundamentally built on human reasoning, empathy, and context-based judgment. When AI begins to play a larger role—especially in critical areas like sentencing, bail decisions, or legal recommendations—there is a real danger that decisions may become overly mechanical or disconnected from the human elements that are essential to fairness. Machines, no matter how advanced, lack the moral reasoning and situational understanding that judges and legal professionals bring to the table.

Compromise of Fair Hearing

Due process ensures that every individual has the right to a fair, impartial hearing. If decisions are made or heavily guided by AI tools—particularly ones that are opaque or difficult to challenge—it may limit the ability of individuals to understand or contest the outcomes. This undermines the core legal principle that everyone deserves to be heard and to receive a reasoned explanation of any judgment against them.

Ethical Responsibility and Accountability

AI lacks ethical consciousness—it cannot be held morally accountable for its decisions. If an AI-driven legal decision causes harm or violates rights, assigning responsibility becomes complex. Unlike a human judge, an algorithm cannot be cross-examined or held accountable in the same way. This raises ethical concerns about fairness, transparency, and who should be liable for AI-driven errors or injustices.

Risk of Over-Reliance on Technology

While AI can support the justice system through data analysis and efficiency improvements, relying on it too heavily may create a false sense of objectivity. AI systems are trained on past data, which may include existing biases or flawed precedents. Uncritically accepting AI recommendations could entrench these issues further, rather than promote justice.

5. Legal Provisions and Regulatory Framework

Information Technology Act, 2000[7]

The Information Technology Act, 2000 serves as the main legislation regulating digital operations and electronic transactions in India. Although the Act does not specifically address artificial intelligence (AI), several of its provisions are relevant to AI-related operations. For instance, Section 43A provides for compensation in cases where personal data is mishandled due to negligence—particularly significant for AI systems that process sensitive user data. Another important clause is Section 72A, which deals with the unauthorized disclosure of personal information obtained through lawful contracts.

[7] IT Act (Information Technology Act, 2000)

Key Case Law: Justice K.S. Puttaswamy and Ors. Vs. Union of India (UOI) and Ors. (AIR 2017 SC 4161)[8]

In this landmark judgment, the Supreme Court of India upheld the right to privacy as a fundamental constitutional right. This decision reinforces the necessity for strict data protection practices, especially in the context of AI technologies that collect and process personal information.

Personal Data Protection Bill, 2019[9]
The Personal Data Protection Bill (PDP Bill), 2019, currently under review, seeks to create a detailed legal framework for handling personal data in India. It outlines critical principles such as obtaining consent, purpose-specific data use, storage localization, and organizational accountability. The bill also calls for the establishment of a Data Protection Authority to monitor compliance and enforce regulations. Importantly, it addresses profiling and automated decision-making, requiring organizations to obtain explicit consent before using AI algorithms that could significantly influence an individual's rights or interests.

Indian Copyright Act, 1957[10]
This Act protects original creative works—such as literary, musical, artistic, and dramatic content—by granting creators exclusive rights over their usage and reproduction. The emergence of AI-generated content has sparked legal debates over ownership and responsibility.

In the *case Gramophone Company of India Ltd. v. Super Cassettes Industries Ltd. (1995) 2 AD 905: (1995) 1 ARBLR 555: (1995) 58 DLT 99: (1995) 33 DRJ 333* [11]

[8] Justice K.S. Puttaswamy and Ors. Vs. Union of India (UOI) and Ors. (AIR 2017 SC 4161)
[9] Personal Data Protection Bill, 2019 (PDP Bill)
[10] Indian Copyright Act, 1957

The Delhi High Court ruled that AI-generated music, being devoid of human creativity, does not qualify for copyright protection. This case helps clarify the legal stance on AI-produced content in India.

National e-Governance Plan (NeGP)[12]

The National e-Governance Plan aims to digitize public services and make government functions more accessible and efficient. AI plays an important role in achieving these goals by automating administrative tasks, supporting data-driven decision-making, and enhancing the quality of services provided to citizens.

New Education Policy (NEP)[13]

India's New Education Policy promotes digital literacy and innovation from an early stage. It introduces coding classes starting from Grade 6, reflecting the government's vision to position India as a global leader in technology and innovation.

AIRAWAT Initiative

NITI Aayog, India's planning body, has launched AIRAWAT—which stands for AI Research, Analytics, and Knowledge Assimilation Platform. This initiative is designed to address the infrastructure, research, and ethical needs related to AI development in India, marking a significant step toward building a robust AI ecosystem in the country.

6. AI in Indian Courtroom Decisions: Relevant Case Laws

As artificial intelligence gradually makes its way into the Indian legal landscape, concerns around its use—particularly regarding privacy, fairness, and accountability—have prompted courts and policymakers to evaluate its

[11] Gramophone Company of India Ltd. Vs. Super Cassettes Industries Ltd. (1995)
[12] National Electronic Governance Plan, 2006 (National e-Governance Plan, NeGP)
[13] New Education Policy (NEP), 2020

implications. While India has not yet seen a definitive case solely addressing the use of AI in judicial decisions, certain landmark judgments lay a foundational framework for regulating and guiding AI applications in law. Among these, Justice K.S. Puttaswamy v. Union of India (2017) is particularly significant.

Indian Supreme Court Cases:

1. Justice K.S. Puttaswamy (Retd.) v. Union of India[14]

Justice K.S. Puttaswamy, a retired judge of the Karnataka High Court, filed a petition challenging the constitutional validity of the Aadhaar scheme launched by the Government of India. The scheme involved collecting biometric and personal data from individuals. The petition raised concerns about the infringement of the right to privacy, especially in the absence of a clear data protection framework.

Key Issue: Recognition of the Right to Privacy as a Fundamental Right
Relevance to AI: In this landmark judgment, the Supreme Court of India unanimously held that the right to privacy is a fundamental right under Article 21 of the Indian Constitution. The ruling has profound implications for AI applications, especially those that involve the collection, processing, and analysis of personal data.

AI systems used in legal and administrative contexts often rely on large datasets, some of which may contain sensitive personal information. This judgment establishes a clear constitutional obligation to safeguard such data and ensures that any AI system used in public decision-making—including in courtrooms—must respect individuals' privacy rights.

[14]AIR 2017 SC 4161.

The ruling also emphasizes informational autonomy, data protection, and consent, all of which are critical when deploying AI in any system that affects personal liberties or legal rights.

2. Anvar P.V. v. P.K. Basheer & Ors (2014) 10 SCC 473[15]

Anvar P.V., a candidate in the Kerala Legislative Assembly elections, alleged that his opponent, P.K. Basheer, had engaged in corrupt practices by making religious appeals to voters through audio CDs. Anvar presented electronic evidence (CDs and digital files) to support his claim, without submitting a proper certificate under Section 65B of the Indian Evidence Act, 1872.

The central legal issue was whether these electronic records could be admitted as evidence without being accompanied by a certificate under Section 65 B (4) of the Indian Evidence Act, 1872. This section mandates that any electronic evidence submitted as secondary evidence must be certified by a person in control of the original electronic device, confirming the authenticity of the content and the integrity of the device used to produce it.

The Supreme Court, in a significant ruling, held that electronic records such as CDs, printouts, or other digital files are not admissible in court unless they are accompanied by the required Section 65B certificate. The Court clarified that merely producing a digital copy is not enough; it must be properly authenticated through statutory procedures. Importantly, the Court **overruled the earlier judgment in *State (NCT of Delhi) v. Navjot Sandhu* (2005)**[16], which had allowed the admissibility of electronic records even without the certificate under certain conditions.

This decision had a far-reaching impact on how digital evidence is treated in Indian courts. It established that strict compliance with procedural

[15] Anvar P.V. vs. P.K. Basheer & Ors (2014)

[16] State (NCT of Delhi) v. Navjot Sandhu (2005)

requirements is necessary to ensure the credibility and admissibility of electronic evidence. The ruling is especially relevant in the age of digital justice and artificial intelligence, as it sets a high standard for the authentication of data generated or processed through electronic or AI-based systems.

In essence, the Court emphasized that technological convenience must not override legal safeguards, and that the use of digital or AI-generated evidence must be held to the same rigorous standards as traditional forms of evidence to protect the integrity of judicial proceedings.

U.S. Supreme Court Case

- *Miranda v. Arizona, 384 U.S. 436 (1966).*[17]

Ernesto Miranda was arrested for kidnapping and rape. He confessed during police interrogation without being told he had the right to remain silent or to have a lawyer present. Then the Issue arose- Can a confession made during police custody be used in court if the suspect wasn’t informed of their constitutional rights?

The U.S. Supreme Court held that the confession was inadmissible. Police must inform suspects of their rights—now called Miranda Rights—before questioning in custody.

European Court of Justice:

Case C-*507/17, Google LLC v. CNIL, ECLI:EU:C:2019:772(ECJ)*[18]

In Case C-507/17, Google LLC v. CNIL (2019), the European Court of Justice was asked to decide whether the “right to be forgotten” under the EU’s General Data Protection Regulation (GDPR) applies globally or only

[17] Miranda v. Arizona, 384 U.S. 436 (1966)

[18]

within the European Union. The case arose when the French data protection authority (CNIL) fined Google for not removing links from all versions of its search engine worldwide, after an individual requested delisting of personal information.

The issue was whether Google was obligated to apply delisting requests beyond the EU domain names, extending to global search results.

The ECJ ruled that the right to be forgotten under EU law does not extend globally. It held that search engines like Google are only required to remove links within the EU domain versions, not from searches conducted outside the EU. However, it also emphasized that search engines must take adequate measures to prevent or seriously discourage users in the EU from accessing delisted content through non-EU domains.

7. AI's Transformation of the Legal System

Traditionally, conducting legal research has been a time-intensive and tedious process, often involving the manual examination of vast legal databases. However, with the rise of AI-based tools, this process has become significantly more efficient, delivering faster results with improved accuracy and relevance.

In India, several advanced AI-powered platforms are transforming legal research:

- Manupatra is a widely used tool that incorporates AI to enhance search capabilities, enabling legal professionals to swiftly locate pertinent case laws and precedents.
- Kanoon.ai, a newer entrant, simplifies the research process by converting complex legal queries into brief, understandable case summaries.

- LegitQuest features a unique iDRAF system—categorizing judgments into Issues, Decisions, Reasoning, and Facts—which allows for quick access to essential case components, streamlining the work of lawyers and judges.

These platforms collectively reduce the manual workload and boost the precision of legal research, allowing practitioners to concentrate more on legal strategy and courtroom arguments.

AI is also revolutionizing legal documentation and contract review, which are traditionally detail-heavy and time-consuming activities. Tools like Kira Systems and ROSS Intelligence employ machine learning to spot errors, flag compliance risks, and suggest edits to enhance clarity and fairness in legal documents.

In the Indian context, startups such as Vakilsearch and LegalKart are applying similar AI technologies to automate contract drafting and document review. These tools not only improve efficiency but also ensure quicker delivery and reduce the likelihood of human mistakes.

Overall, AI is reshaping the legal profession by automating routine tasks, improving research quality, and enhancing the accuracy of legal documentation.

8. Augmented Reality (AR) and AI in Courtrooms

Augmented Reality (AR), which overlays digital information onto the real world, has expanded beyond its original use in gaming and now spans multiple industries, including retail, healthcare, education, and entertainment. Its ability to seamlessly integrate virtual elements into real-world environments offers users an enhanced and interactive experience. In the legal field, AR is increasingly being used in courtrooms, revolutionizing how evidence is presented and understood.

By improving how jurors perceive and engage with evidence, AR is transforming traditional judicial processes and offering new ways to approach complex cases. This technology not only aids in enhancing juror comprehension but also boosts efficiency in legal proceedings. In practice, AR enables attorneys to present case details, such as crime scene reconstructions or anatomical diagrams, in a more digestible and visually engaging manner, helping jurors understand critical aspects of a case more clearly.

Moreover, the use of AR in courtrooms can streamline various stages of a trial, improving case presentations and reducing the time needed to convey key information. Legal professionals can leverage AR to simulate events, making it easier for participants to process complex details during trials.

- Ethical and legal concerns about deepfake evidence and AI-generated testimonies

o Deepfakes and AI-generated testimonies raise concerns about the authenticity of evidence. A deepfake can create hyper-realistic images, videos, or audio clips that appear to be genuine but are entirely fabricated. This undermines the reliability of evidence in court, making it challenging for judges, lawyers, and juries to distinguish between real and manipulated content. If these AI-generated materials are introduced as evidence, they could easily mislead the court, potentially resulting in wrongful convictions or acquittals.

o Using deepfake technology to create false testimonies or manipulate video or audio recordings can be used unethically to distort facts. If used maliciously, deepfakes could fabricate confessions, testimonies, or witness statements, undermining the integrity of legal proceedings. This raises serious concerns about justice, fairness, and the possibility of wrongful convictions based on falsified or misleading information.

9. Conclusions and Recommendations

Conclusions

The integration of Artificial Intelligence (AI) and deepfake technologies into the legal system presents both significant opportunities and challenges. While AI has the potential to enhance legal processes—by improving efficiency in legal research, streamlining trial phases, and aiding in decision-making—there are serious ethical, legal, and security concerns that must be addressed. The emergence of AI-generated testimonies and deepfake evidence threatens to undermine the authenticity of evidence and judicial fairness, leading to the possibility of wrongful convictions or acquittals. The lack of clear guidelines on the admissibility of such evidence and the challenges surrounding the accountability for AI-generated materials pose serious risks to the integrity of the legal system.

Moreover, as AI technologies continue to evolve, the implications for privacy, human rights, and due process must be carefully considered. The potential for deepfakes to manipulate visual or audio content in misleading ways could erode public trust in the judicial system, making it essential for stakeholders to proactively address these issues.

Recommendations

1. Establish Clear Legal Guidelines

Legal systems must develop and adopt clear rules and frameworks to address the admissibility of AI-generated evidence. These guidelines should include standards for verifying the authenticity of digital content and provide protocols for identifying deepfake material before it is allowed in court. Courts should also incorporate mechanisms for cross-referencing and validating digital evidence to prevent manipulation.

2. **Create Ethical Standards for AI Usage**

Lawmakers and legal professionals should collaborate to establish ethical guidelines for the use of AI in legal practice. These standards should ensure AI tools are used to complement human judgment rather than replace it, particularly in sensitive areas such as witness testimonies or judicial decision-making. AI should be transparent, explainable, and not compromise fundamental human rights.

3. **Introduce AI Literacy Training for Legal Professionals**

Legal professionals, including judges, lawyers, and court staff, should receive training on AI technologies and their implications for legal practice. Understanding AI's potential and limitations will enable them to better navigate the complexities of using AI-driven tools, spot deepfake evidence, and make informed decisions about the role of AI in the courtroom.

4. **Implement Robust Data Security Measures**

Given the potential for AI systems to be vulnerable to attacks or misuse, it is crucial to implement strong data protection and cybersecurity protocols within legal AI systems. This will ensure the integrity of AI-driven legal tools and prevent unauthorized access to sensitive case data.

5. **Develop an Independent Oversight Body**

Establishing an independent regulatory body or authority to monitor the use of AI in legal proceedings could help ensure compliance with ethical and legal standards. This body could oversee the development and deployment of AI tools in legal practice, ensuring that they are used responsibly and fairly.

6. **Strengthen Privacy Protections**

With the rise of deepfake technology, protecting individuals' privacy is increasingly critical. Legal systems should strengthen privacy protections to ensure that AI-generated content does not violate the rights of individuals. Additionally, strict consent processes should be enforced when using AI systems to process sensitive personal data.

7. Promote Public Awareness and Trust

Public education campaigns should be initiated to inform citizens about the role of AI in the legal system, including the risks associated with deepfake evidence and AI-generated testimonies. Building trust and understanding among the public is crucial to ensuring confidence in the judicial system, particularly as these new technologies become more integrated into legal practices.

10. References

- B.V.Seshagiri Advocate. "'5 Crore Cases': India's Judiciary Collapsing under Weight of Huge Backlog." *Newsmeter.in*, NewsMeter, 2 Dec. 2024, newsmeter.in/interviews/5-crore-cases-indias-judiciary-collapsing-under-weight-of-huge-backlog-739920?utm_. Accessed 15 June 2025.

- Coglianese, Cary, et al. "AI in the Courts: How Worried Should We Be? | Judicature." *Judicature.duke.edu*, 6 Mar. 2024, judicature.duke.edu/articles/ai-in-the-courts-how-worried-should-we-be/.
- "Indian Courts Achieve Milestone in Case Disposals – GKToday." *Gktoday.in*, 2025, www.gktoday.in/indian-courts-achieve-milestone-in-case-disposals/?utm_. Accessed 15 June 2025.

- Khan, Aamir. "AI-Powered Indian Judiciary: A Step Forward or Cause for Concern?" *Bar and Bench - Indian Legal News*, 6 June 2023, www.barandbench.com/columns/litigation-columns/ai-powered-indian-judiciary-a-step-forward-cause-concern. Accessed 15 June 2025.

- Luminance Technologies Ltd. “Top 20 Global Law Firm Adopts Luminance’s AI for Large Scale Contract Review.” *Luminance.com*, 2022, www.luminance.com/news/press/20220524_top_20.html?utm_. Accessed 15 June 2025.

- PPJ. “NJDG-National Judicial Data Grid.” *Ecourts.gov.in*, 2024, njdg.ecourts.gov.in/njdg_v3/.
- Rocha, Cinara, and João Alvaro Carvalho. “Artificial Intelligence in the Judiciary: Uses and Threats.” *ResearchGate*, unknown, 4 Dec. 2023, www.researchgate.net/publication/376191270_Artificial_Intelligence_in_the_Judiciary_Uses_and_Threats.

- The Precedent. “AI and the Future of Justice: How Technology Is Reshaping India’s Legal Landscape in 2025.” *The Precedent*, 13 Feb. 2025, precedent.in/artificial-intelligence/ai-future-justice-how-technology-reshaping-india-legal-landscape/. Accessed 15 June 2025.

EXTRADITION LAW: ROLE OF AI IN INTERNATIONAL LEGAL DISPUTES

Bhaggya s. Thakre and Anjali wani[1]

Abstract

Artificial Intelligence (AI) is increasingly transforming legal systems, including the field of extradition law, which governs the transfer of individuals between jurisdictions for legal proceedings or sentence execution. AI technologies, such as predictive analytics and automated case management systems, enhance efficiency by streamlining workflows, managing large volumes of legal data, and supporting risk assessments. These tools are particularly beneficial in complex, cross-border extradition cases. However, the integration of AI also raises significant legal and ethical concerns, including algorithmic bias, data privacy, and the need for human oversight. Case studies reveal both the benefits and pitfalls of AI in this context, underscoring the importance of international cooperation and regulatory reform. Moving forward, a balanced approach—blending technological innovation with strong ethical standards and global collaboration—will be crucial to leveraging AI's full potential in advancing justice within extradition law.

1. Introduction

Artificial Intelligence (AI) is fast changing various fields, including law enforcement and the criminal justice system. As AI technologies have the potential to increase efficiency and effectiveness, their application in policing has raised essential questions regarding privacy, the right to a fair trial,

[1] Students of BBA LLB 2nd Year

equality, and non-discrimination. Several nations are starting to formulate regulatory frameworks to counter these threats and ensure that AI applications do not contravene basic human rights. This research aims to explore the existing state of AI regulation in European security and criminal justice systems. Through the analysis of various regulatory strategies, the report will determine commonalities and particular legal actions that have been taken. It will also present real-life examples demonstrating how these regulations are enforced, providing a better understanding of AI governance in the justice system.

The growing dependence on AI technologies in criminal justice has necessitated regulatory control to ensure human rights and legal safeguards. AI can contribute significantly to enhancing cross-border judicial cooperation and simplifying justice-related processes. The judicial system digitalization has precipitated the embrace of AI-centered solutions that not only increase judicial authorities' effectiveness but also enhance access to justice and lower legal proceedings time eventually. The breakthroughs, on the other hand, come along with ethical, as well as legal, obstacles that require wide-ranging policies developed to avoid improper use and predisposition in applying AI. Governments and institutions need to balance maximizing the potential of AI with retaining safeguards that ensure the fundamental rights of individuals.

2. Comprehensive overview of extradition law and it's implications.

Extradition refers to a legal process by which one state hands over an accused or convicted person to another for prosecution or punishment. It guarantees that criminals cannot evade justice by fleeing across borders and is regulated by bilateral or multilateral treaties specifying crimes, procedures, and conditions. Although extradition encourages cooperation among nations, it can put national sovereignty to test and be in conflict with human rights. Most countries will not extradite if there is a risk of unfair trials, persecution, or the death penalty. Moreover, extradition decisions affect diplomatic relations, particularly in sensational cases, making it an important and challenging feature of global criminal justice.

2.1 Historical Context

Extradition is regulated domestically by extradition acts and internationally by diplomatic treaties. Belgium pioneered an extradition law in 1833, concurrent with the first asylum law.

Extradition acts prescribe extraditable offenses, set out procedures, and regulate the interaction between domestic legislation and international agreements. Developments in extradition law have influenced contemporary French and American practice. This article explores the historical evolution of extradition legislation and its effect on current legal systems of both countries. Although extradition is not always treaty-based, extraditions based on treaties have been in existence since history began. The oldest known diplomatic treaty has provision for extraditing people for ordinary crimes and not political crimes. The historical study of extradition enlightens us to the way contemporary extradition regimes developed, attempting to reconcile national legal systems with transnational compacts.

2.2 Key Principles of Extradition

The **double criminality, or dual criminality**, is a basic principle of international extradition law. It requires that a crime for which extradition is requested should be a crime in both the requesting and requested states. Double criminality ensures that a person is not extradited for conduct that is not criminal in the requested state, thus ensuring justice and preventing unfair prosecutions. It also provides a threshold of gravity for extradition cases so that small or non- criminal offenses will not result in extradition.

The **norm of reciprocity, or the principle of reciprocity**, is an inherent principle of international law that encourages mutual respect, cooperation, and equality among nations. It makes sure that nations provide each other with reciprocal legal obligations, especially in the area of extradition. Under this principle, states commit themselves to extradite individuals who are accused or convicted of crime, on condition that the requesting state also shows reciprocal cooperation. This promotes fairness and balance in international relations by holding states accountable for their legal obligations.

The **principle of specificity** limits the prosecuting power of the requesting state. It provides that a person extradited for a specific crime can only be tried for that exact crime and not for any other earlier crime. Secondly, the extradited person cannot be sent to a third nation for any crime committed prior to extradition except when an exception exists or the extradite voluntarily waives this right. This principle safeguards the right of asylum and discourages states from abusing extradition as a means to pursue individuals for political or extraneous crimes.

The **aut dedere aut judicare** principle creates an international law obligation on states to prosecute those suspected of committing serious international crimes or extradite them to another state that desires to prosecute them. This principle is extensively integrated into global treaties concerning international crimes, making sure that the perpetrators are brought to justice. This principle, however, is not considered a rule of customary international law as shown in the Lockerbie case. It applies irrespective of either the accused or the victim's nationality and is still an important instrument against international crimes.
Collectively, these principles underpin extradition law so that the cooperation of states in the application of the law is done in a fair, due process manner, and according to international norms of law.

2.3 International Treaties and Agreements

The European Parliament's 6 October 2021 resolution considers the application of artificial intelligence (AI) in criminal law, highlighting both its advantages and dangers. Although AI has the potential to advance law enforcement through the identification of crimes such as financial fraud and cybercrime, it also threatens through bias, discrimination, privacy breaches, and false arrests. The resolution emphasizes that AI should respect basic rights, guaranteeing transparency, accountability, and fairness. It cautions against mass surveillance through AI and demands that human oversight is indispensable. AI cannot substitute human judgment in criminal justice or erode the presumption of innocence and fair trial rights.[1]

AI technology is also subject to cyberattacks and data tampering, which may undermine legal processes. To mitigate these threats, the EU demands a robust legal framework that governs the use of AI in law enforcement, making its application ethical, justified, and proportionate. AI must not result in discrimination, mass surveillance, or unjustified profiling. Tight regulation is necessary in order to avoid AI from perpetuating bias in policing and judicial rulings. AI choices should always be open to human intervention, and legal accountability must always be with an individual. Security threats, such as data breaches, need to be addressed, and AI technologies have to meet ethical and legal guidelines. The resolution also demands independent assessment and regular audits of AI in law enforcement. AI systems should be transparent, with authorities revealing their deployment. The EU demands bans on biometric mass surveillance, AI-driven social scoring, and private surveillance databases such as Clearview AI. Ethical AI training must be given to law enforcement, and predictive policing must not be the exclusive method.

Lastly, the EU promotes global cooperation to create moral AI standards to ensure that AI technologies in law enforcement uphold universal human rights.

3. The role of AI in Legal Framework

Artificial Intelligence (AI) is revolutionizing the legal profession with enhanced efficiency, accuracy, and decision-making. AI software tools automate document examination, saving time and minimizing errors in contract handling and due diligence. Predictive analytics enable attorneys to predict the outcome of cases and create solid litigation strategies. AI- powered legal research accelerates information retrieval through the analysis of massive databases of legal documents, case laws, and precedents. In legal proceedings, AI helps lawyers spot applicable precedents and visualize patterns in case results, supporting planning.

AI also promotes access to justice by lowering litigation expenses and simplifying legal procedures. Challenges still exist, such as the necessity for regulating frameworks to accommodate AI innovation and ethical issues with respect to its use in judicial decision- making. While AI can never substitute human judgment, it is transforming legal practice ([1]*European Parliament Resolution on AI in Criminal Law, 6 October 2021)*through increased productivity and strategic thinking. Its adoption should be balanced by rules to ensure equity and responsibility in the justice system.

3.1. AI Technologies in Legal Practices

The Council of Europe adopted the world's first international treaty to control AI in a manner that safeguards human rights, democracy, and the rule of law. Available for signature by non- European nations, the treaty enforces legal principles on the full life cycle of AI, keeping innovation balanced against risk. Taking a risk management approach, the treaty necessitates a meticulous weighing of the negative consequences of AI.

Implemented in Strasbourg at the Council of Europe's annual ministerial session, the Framework Convention on Artificial Intelligence is the culmination of two years of negotiation among 46 European member states, the EU, and 11 non-member states, together with private sector and civil society actors. The treaty covers both public and private sector AI applications, with the obligation of parties to create transparency, oversight, and accountability. It obliges risk assessments, protection against discrimination, privacy infringements, and dangers to democracy, and guarantees legal recourse for AI-related human rights abuses. Governments are also obligated to prevent AI from compromising democratic institutions, judicial independence, or the separation of powers. *(Framework Convention on Artificial Intelligence Council of Europe, 2024)*

This treaty represents a worldwide push to control AI in a responsible manner, such that its gains do not happen at the expense of basic rights and freedoms. The primary methods by which legal professionals are employing generative AI within their practice are:

Drafting/templating messages (e.g., memos, emails, letters to opposing counsel, etc.): 58% Searching the law: 53%

Summary of legal narratives: 42% Reviewing legal documents: 34%

Drafting/templating legal agreements: 23% Due diligence: 21%

Reviewing discovery: 15%

Negotiating/redlining contracts: 11%

Preparing case filings (e.g., pleadings, motions, jury instructions, etc.): 8% Estate planning: 2%

4. AI Applications in Extradition Process

AI is increasingly being used to support extradition procedures, primarily through data analysis, document processing, and decision support. It doesn't make decisions regarding extradition, but it assists in the collection and analysis of data, suspect profiling, and legal document management.

AI tools have the capability to process large datasets to evaluate risks such as flight or reoffending and assist in the identification of extradition-subject individuals. Translation and document processing are automated, enhancing legal processes and cross-border cases. AI assists judges with case law, precedents, and risk analysis, enhancing decision-making. Predictive analytics can prioritize cases by approximating extradition success rates. AI usage in extradition does, however, pose issues concerning bias, privacy, and human rights. Transparence and ethical guidelines must be followed to avoid unjust conclusions. Global coordination is required for AI applications to be harmonized with legal and ethical standards.

AI can optimize extradition to become more efficient with fairness and conformity if managed suitably.

5. CHALLENGES OF IMPLEMENTING AI IN EXTRADITION LAW

Using AI in extradition law comes with some tough hurdles. It's crucial to make sure the process stays fair and doesn't let any bias creep into the algorithms. Protecting sensitive data and keeping it secure is another big concern. On top of that, there needs to be clear accountability for decisions made by AI systems. Finally, legal frameworks must evolve to keep pace with these advanced technologies. It's a balancing act to get it all right!

Challenges in the process of extradition law:-

1. Legal and ethical concerns
2. Technical challenges
3. Legal and regulatory challenges

1. **legal and ethical concerns**

Using AI in extradition law comes with several challenges that need careful attention. For starters, AI can carry existing biases from the data it learns, which might lead to unfair treatment of certain groups. Then there's the issue of transparency—some AI systems work like a "black box," making decisions that are hard to explain, which can weaken trust in the legal system. Accountability is another tricky area; when AI makes mistakes or causes harm, it's not always clear who's responsible. Privacy is a big concern too, since extradition cases involve sensitive personal data that must be protected and handled according to strict rules.

Most importantly, AI should never undermine basic human rights, and decisions involving AI must always follow fair legal processes. It's a complex balancing act to make sure AI is used responsibly in such critical matters.

2. **Technical challenges**

Using AI in extradition processes comes with several hurdles that need careful navigation. The reliability of decisions made by AI depends heavily on the quality of the data used to train it—if the data is flawed, so are the results. Legal reasoning is another tricky area, as AI often struggles with the subtle nuances and complexities that are key to fair and accurate outcomes. On top of that, integrating AI into existing legal systems and workflows isn't always smooth sailing, and there can be technical and operational challenges. Cybersecurity is also a big concern, as AI systems are at risk of being hacked, which could jeopardize the sensitive data and processes involved. Finally, implementing and maintaining AI systems demands a workforce with specialized skills, and there might not always be enough qualified professionals in the legal sector to handle this. Each of these challenges requires careful planning and oversight to ensure AI is used responsibly.

3. legal and regulatory challenges

The use of AI in extradition processes brings forward several pressing issues that demand attention. Current legal frameworks often struggle to keep up with the unique challenges posed by AI, requiring significant adaptation to ensure they remain relevant. Since extradition involves cross-border cooperation, it is essential to foster international collaboration to ensure AI is used responsibly and ethically in this field. A major hurdle is the absence of clear regulations and guidelines governing the use of AI in law enforcement and extradition, which creates uncertainties. Accountability is another key aspect; it's crucial to establish clear responsibilities for AI-driven decisions to maintain fairness and transparency. Additionally, the ever-changing threat landscape poses risks, and AI systems need to be flexible enough to adapt to these evolving challenges. Addressing these concerns is vital for the ethical and effective use of AI in extradition.

6. RECENT CASES OF EXTRADITION

Vijay Mallaya's case[2]

The case of Mr. Vijay Mallaya, the business tycoon and owner of Kingfisher Airlines and United Breweries Holdings Ltd., is arguably the most well-known extradition case in India He owed a whopping debt of over ₹6,000 crores to 17 Indian banks including the State Bank of India and the Indian Overseas Bank. Fearing an impending arrest, Mallaya fled from India to the United Kingdom in 2016. His extradition was sought by India in 2017.Mallya's extradition case was laid before the Westminster Magistrate's Court in London. In 2018, the Court ordered his extradition to India. His appeal at the High Court in London was rejected; however, he has not been brought back to India yet due to ongoing legal procedures.

It's also worth noting that in 2019, he was declared a 'Fugitive Economic

[2] Dr Vijay Mallya v. State Bank of India, (2018), Westminster Magistrate's Court, London

Offender' under the Fugitive Economic Offenders Act, 2018.[3]

Nirav Modi's case:[4]

Mr Nirav Modi was a luxury diamond jewellery merchant. In 2018, the Punjab National Bank (PNB) filed a complaint before the Central Bureau of Investigation (CBI), alleging Nirav, along with his wife Mrs Ami Modi, of fraudulently obtaining fake Letters of Understanding (LoU) worth ₹11,400 crores. The money was then channelised to his fifteen overseas sham companies.

Following a CBI probe, the Enforcement Directorate (ED) confiscated Nirav's assets in India. He fled India and sought asylum in the United Kingdom. Interpol issued a Red Corner Notice against him in 2018. Following an extradition request from India, a Westminster Court issued an arrest warrant against Nirav. The Court ordered his extradition to India in 2021.[5]

The Mehul Choksi Case:

The case involving Mr. Choksi was one that stirred great controversy in India. Mr Choksi is wanted in India for counts of criminal conspiracy, corruption, money laundering and criminal breach of trust on account of the Punjab National Bank Loan Fraud.

After being accused of his crimes, Mr. Choksi fled to Antigua, where he purchased citizenship under am investor scheme, in light of avoiding deportation to India for his Trial.

The Antiguan authorities are naturally reluctant to extradite one of their citizens as they believed he would be subject to inhumane conditions in India. Thus, India is facing a difficult time retrieving Mr. Choksi from Antigua.

[3] Fugitive Economic Offenders Act, 2018 (India).

[4] Punjab National Bank v. Nirav Modi & Ami Modi, 2018, Central Bureau of Investigation (CBI) Complaint

[5] Interpol Red Corner Notice, 2018

The Ravi Pujari Case:

Ravi Pujari was a popular gangster in the early 2000s who was known for threatening eminent personalities in the film and real estate industries. He was wanted by officials on counts of murder of Mr. Kukreja, a popular builder and on attempt to murder charges by Mr. Suresh Wadhwa. Mr Pujari fled India and remained a fugitive in numerous countries such as Australia, The United Arab Emirates, Burkina Faso and Senegal. His threat calls to a Kerala MLA traced his location to Senegal, where he lived under the alias of Anthony Fernandez.

The fact that India had extradition arrangements with Senegal allowed Mr Pujari to be extradited to Bangalore in the subsequent days.

WIKILEAKS FOUNDER JULIAN ASSANGE[6]

Wikileaks founder Julian Assange has been fighting extradition to the United States on

espionage charges since he was arrested in London in 2019, after spending seven years there holed up in the Ecuadorian Embassy.

Assange, an Australian citizen, was charged by U.S. federal prosecutors with publishing secret diplomatic cables and military reports on his Wikileaks site, in what U.S. authorities have dubbed one of the biggest leaks of classified information in history.

A UK court initially blocked Assange’s extradition due to concerns over his mental health, but the United States successfully appealed the decision.

Assange appealed in August. He has said he is being persecuted for his political beliefs and that he was acting as a journalist in publishing the leaked documents.

[6] Julian Assange v. United States, 2019.

Michael Taylor and Peter Taylor[7]

Michael Taylor and his son, Peter, were extradited from the United States to Japan in March 2021 for helping former Nissan Motor Co Ltd. Chairman Carlos Ghosn flee Tokyo in 2019 after being charged with financial crimes.

Japanese authorities said the Taylors hid Ghosn in an audio equipment box and smuggled him onto a plane to his native Lebanon which has no extradition treaty with Japan.

The Taylors asked a federal court to block their extradition to Japan, saying "bail jumping" is not a crime in Japan and that they would be subjected to “mental and physical torture” if

incarcerated there.

A federal judge rejected the Taylors' petition and cleared their extradition in January 2021. The U.S. Supreme Court rejected their appeal of the decision two months later.

CASE STUDIES OF AI IN EXTRADITION

Successful Implementations:

AI has been used effectively in justice systems in various countries. For example, in India,

tools like SUPACE have helped make court processes smoother by speeding up research and improving efficiency. In France, AI has shown potential in handling civil and criminal matters, helping to streamline decision-making and reduce case backlogs.

Failures and Lessons Learned:

Not all AI projects have been successful. For instance, IBM Watson for Oncology faced issues because it relied on artificial data, leading to inaccurate results.[8] Similarly, Amazon’s

[7] Michael Taylor & Peter Taylor Extradition Case, 2021.

[8] IBM Watson for Oncology

AI recruitment tool showed bias against women, reminding us how critical it is to use diverse

and high-quality data.[9] These cases highlight the need for ethical oversight and thorough testing before deploying AI.

Comparative Analysis of Jurisdictions:

Extradition rules differ across countries, affecting how AI can be integrated. For example, the

U.S. requires treaties for extradition, while countries like Germany and Switzerland can operate on reciprocity without formal agreements.[10] Nations like India, the USA, and European countries approach extradition laws uniquely, reflecting their distinct legal and political landscapes.

7. FUTURE OF AI IN LAW

The idea of using Artificial Intelligence (AI) in the legal system is quite impressive! During the pandemic, courts struggled as in-person hearings stopped, leading to delays in justice. To tackle this, the Supreme Court of India introduced **virtual hearings** in March 2020, where

judges worked from home. Other courts, like the Bombay and Patna High Courts, also shifted to virtual hearings when COVID cases rose. This shows how the judiciary quickly adapted to modern solutions for unique challenges.

AI is now being explored to improve the legal system even further. For example, the Supreme Court launched **SUPACE** (Supreme Court Portal for Assistance in Court Efficiency)[11], which helps automate parts of the judicial process to save time and reduce delays. Another initiative, **SUVAS**[12],

[9] Amazon AI Hiring Tool, 2018 — Discontinued for gender bias, showing importance of diverse data

[10] United States Extradition Treaties Database.

[11] AI in Germany and Switzerland

[12] SUPACE (Supreme Court Portal for Assistance in Court Efficiency)

translates judgments into regional languages like Marathi, Hindi, Tamil, and more. This makes justice more accessible to people who don't speak English.

AI works by analyzing huge amounts of data to find patterns and make predictions. It's already being applied in areas like criminology, law, and forensics. AI can assist with

decisions on bail, parole, and sentencing by identifying patterns in past cases. However,

experts worry that AI might sometimes be biased because it relies on historical data, which can affect its fairness.

AI can also support lawyers by helping with legal research, which is essential for building strong cases. Tasks like analyzing laws, reasoning, and preparing arguments can become much easier with AI.

To sum up, AI can make court proceedings faster, more efficient, and accessible, benefiting both judges and lawyers. While there are challenges, such as the risk of bias, it's important to embrace this technology to modernize the judiciary. After all, progress is necessary, and fear of new tools might hold us back. Every innovation has its pros and cons, but the potential benefits of AI in the legal system make it worth exploring!

- **EMERGING TECHNOLOGIES**

Modern problems need modern solutions, and the rise of AI is a perfect example of this. AI has brought about new challenges in areas like Intellectual Property Rights (IPR) and Human Rights, which now need updated laws to address these issues. For instance, AI robots have become so advanced that they can create or invent things on their own. This raises some important questions: If an AI creates something, who gets the credit? Will the patent belong to the robot, or to the person who created the robot? These are tricky questions that need answers.

Traditionally, humans have copyrights, patents, and trademarks to protect their inventions. But with AI entering the picture, we might even need to think about creating new laws or rights specifically for robots. Robots don’t have personal freedoms like humans, but their growing intelligence could give them a reason to demand their own set of laws and rights.

This issue became even more important after Saudi Arabia granted citizenship to Sophia, the world’s first social humanoid robot. This move inspired other countries like China and North Korea to consider doing the same. Even Indian scientists are working to develop advanced humanoid robots like Sophia. Clearly, this is a growing trend, and it’s essential for our legal systems to start preparing for these emerging challenges before it’s too late.[13]

- **INTERNATIONAL COOPERATION AND STANDARDS**

International Model Laws on extradition[14]

The Geneva Conventions and their Additional Protocols (1949) were some of the earliest

conventions that dealt with extradition to some extent; it recognised the state's cooperation in extradition. Thereafter, most countries have signed several multilateral and bilateral treaties on extradition. For instance, the United States of America has signed extradition treaties with over 100 countries. Various countries have also incorporated provisions for extradition in their penal codes.

The United Nations Model Treaty on Extradition (1990)[15]

The UN Model Treaty on Extradition firmly emphasised international cooperation in extradition-related matters. It has 18 Articles, dealing with the grounds for refusal of extradition requests, Rule of Speciality, etc. However, it prioritises the discretion of the territorial State.

[13] Saudi Arabia grants citizenship to Sophia the robot, 2017.

[14] Geneva Conventions and Additional Protocols, 1949.

[15] The United Nations Model Treaty on Extradition, 1990.

The United Nations Model Law on Extradition (2004)[16]

The UN Model Law on Extradition is inspired by the UN Model Treaty and aims to enhance international cooperation in extraditions. It also aims to act as a supplementary statute in cases of countries where extradition treaties are absent. Sections 5 and 6 of the Model Law explicitly provide that extradition shall not be granted if, in the view of the territorial State,
the extradition is requested for torturing or punishing the fugitive on the basis of his caste, ethnic origin, race, etc.

8. CONCLUSION

Artificial intelligence can make the legal system faster, fairer, and easier for people to access. For example, it can speed up decisions, analyze legal factors more thoroughly, and help more people get justice. But it's also important to address the ethical challenges of using AI in the judicial process. To use AI responsibly in law, we need to consider issues like transparency, fairness, data privacy, and security. AI systems should produce results that are easy to understand, so legal professionals and the public can see how decisions are made. To prevent bias and ensure fairness, we need to carefully select training data, regularly check AI systems for problems, and collaborate across different fields to challenge inequality. Protecting people's privacy is critical. Strong security measures, like anonymizing personal information and getting consent before using data, are essential to avoid data leaks or misuse.

AI shouldn't replace human judgment entirely—it should assist legal professionals while keeping human oversight in place. Clear ethical rules and legal regulations need to be established and updated regularly to guide the responsible use of AI. By following these principles, we can make the most of AI's benefits while minimizing risks. A thoughtful approach will allow AI to

[16] The United Nations Model Law on Extradition, 2004.

contribute to a legal system that is faster, fairer, and more accessible to everyone. Lawyers, lawmakers, and other stakeholders have an important role in ensuring AI is used ethically, with respect for human dignity and the rule of law.

References

- Council of Europe. (2024). Framework Convention on Artificial Intelligence. Strasbourg.
- European Parliament. (2021). Resolution on Artificial Intelligence in Criminal Law. Brussels.
- Government of India. (2018). Fugitive Economic Offenders Act. Ministry of Law and Justice.
- IBM Corporation. (2019). Watson for Oncology: Lessons Learned. IBM Research.
- Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.
- Plachta, M. (2001). Extradition and Human Rights: Selected Issues. Netherlands International Law Review.
- Shearer, I. A. (1971). Extradition in International Law. Manchester University Press.
- Slobogin, C. (2019). Predictive Policing and Artificial Intelligence. Yale Law Journal, 128(5), 1305–1342.
- Surden, H. (2014). Machine Learning and Law. Washington Law Review, 89(1), 87– 115.
- United Nations. (1990). Model Treaty on Extradition. New York: United Nations Office on Drugs and Crime.
- United Nations. (2004). Model Law on Extradition. Vienna: UNODC.

DEEPFAKES: BLESSING OR CURSE IN MODERN TECHNOLOGY?

Ananya Srivastava and Upashna Sharma[1]

Abstract

Deepfakes have emerged as one of the most fascinating but troubling advancements driven by artificial intelligence (AI) and deep learning in an era where visual content dominates communication. The potential to create hyper-realistic synthetic movies, audios, and images that are indistinguishable from authentic content has called into question the entire foundation of confidence in digital media. Deepfakes are a combination of deep learning and fake referring to use of AI, machine learning, algorithms, also Generative adversarial network are a type of machine learning model consisting of two neural networks — A generator and A Discriminator that work against each other in a competitive process. As the generator creates fake content and the Discriminator detects fake content. Deep fakes technology which was originally intended for amusement, today drives misinformation, manipulation, fraud, increased sophistication makes it difficult to discriminate between legitimate and fraudulent material, weakening media confidence. With The Effects of Realism and Excellent Quality, Deepfakes Poses a Significant Issue and It Makes It Harder to Differentiate Between Authentic and Modified Information. For Conducting this research my motivation is to highlight The Increasing Danger Which Deepfakes Are Causing in Society, Especially in Terms of Fraud, Misrepresentation and Privacy Invasion. All-Platform Universal Detection Tool's Absence Which Can Consistently Detect Deepfakes is the main research gap founded by me. There Is Not Yet Any AI-Driven Detection

[1] BBA LL.B 2nd Year Law Students
Bharati Vidyapeeth (Deemed to be University), New Law College, Pune

Solution Which Can Offer 100% Accuracy, which is an Ongoing Problem Nowadays. Thus, More Flexible Methods Development is essential halt Deepfakes rapid advancement and digital authenticity. This study aims to Address Deepfakes Issues and to find Out solutions To Enhance Flexibility Pertaining to The Identification of Deepfakes. This Study Has Comprehensive Detection Framework to integrate social media And Forensic deepfake videos with benchmark datasets such as Face Forensics++ And Celeb-DF. While facial speech disintegration is being processed by CNN's and Transformers, pixel distortion, voice modification, and micro expression forensic and biometric analysis look for other anomalies. The machine learning model is developed to counter newly emerging deepfake techniques. Upon achieving accuracy through cross-dataset testing, the deepfake detection framework is deployed into cybersecurity, media verification, and digital forensics. Furthermore, this examines the ethical perspective of AI by trying to ensure a balance between privacy and detection while proposing regulatory strategies for digital integrity assurance. The Deepfake Detection Model performance varies depending on the models and dataset used. CNNS' (Convolutional neural networks) and transform based architectures have achieved accuracy rates between 85% and 98% and these are recent based AI models. Nevertheless, testing on real-world diverse datasets reveals reduced performance due to differences in lighting, resolution, and adversarial attacks. One of the best approaches, frequency-based analysis, has been most effective in identifying real and spurious videos. But as GANs evolve, synthetic media becomes more convincing, and detection becomes more challenging. Detection tools are promising, but they will have to keep pace with the constant advancements in Al-generated content. Ethical issues and legal structures will have an important role to play in guiding the ethical application of deepfake technology in the future.

Introduction

In The 21st Century, The Era of Technology Has Been Growing Rapidly in Digital Media, Which Has Led to the Evolution of Deepfake Technology. Deepfake Technology Is Emerging as Both Revolutionary Innovation as Well as Potential Societal Threat. Deepfakes Are Created by Using Artificial Intelligence and Deep Learning Techniques, allowing for the virtual Manipulation of Images, Videos and Audio. Many Celebrities and High-Profile People Have Become Victim of This Deepfake Technology's Misuse. This Technology Makes It Difficult to Distinguish Fabricated Content from Reality. As These Innovations May Help in Creativity and Educational Applications, They Also Raise Major Serious Concerns Related to Misinformation, Ethical Dilemmas and Identity Fraud. As Deepfake Technology Continues to Progress, it is very crucial to know and analyze its implications, regulatory challenges, ethical implications, detection strategies, Potential threats And Future Developments, which could lead to cyberbullying in this digital age. This Paper Explores the Origins, Applications, Risks, Threats and Solutions Related to Deepfakes, Highlighting Their Impact on Digital Media and Public Trust.

Methods and Materials:

Deepfake Detection and Analysis Materials: This study brought together publicly available datasets, open-source tools, and AI-driven platforms to explore the creation and forensic analysis of deepfakes. Among the datasets used are: A key resource for training and testing deepfake detection models, Face Forensics++ serves as a widely recognized benchmark dataset that includes both real and altered videos employing various facial forgery techniques. To reflect the real-world diversity in facial expressions, head poses, and backgrounds, Celeb-DF (v2) features high-quality deepfake videos of celebrities. The Deepfake Detection Challenge (DFDC) dataset, released by Facebook AI, includes a variety of deepfake content specifically selected to evaluate model robustness. To capture multimodal fake content, additional materials were sourced from synthetic speech databases and GAN-generated video archives. Comparative data analysis and statistical validation were conducted using MATLAB and SPSS software tools. For supervised learning, frames were labeled and assessed with annotation tools like Label box and VGG Image Annotator (VIA). To facilitate scalable model deployment, the research was performed on cloud-based systems such as Google Collab Pro+ and Kaggle Kernels.

Materials and Methods

The approach to study the technological and legal aspects of deepfake detection involved a multidisciplinary research approach pairing computational experiments with legal doctrinal analysis. The computational aspect of the research examined deepfake video detection using convolutional neural network (CNN) architectures. In particular, the XceptionNet and EfficientNet CNN architectures were specifically used due to their demonstrated effectiveness in video forensics.[1] The models were trained and tested using three publicly available benchmark datasets: FaceForensics++, Celeb-DF v2, and Deepfake Detection Challenge (DFDC) Preview Dataset.[2] Each of the datasets contained a corpus of both real, and synthetically manipulated, facially video footage and included natural variation and manipulation approaches.

Prior to feeding video to the models, a series of preprocessing tasks were performed on each of the videos. The tasks included aligning faces, resizing frames to an input rating of 256✕256 pixels, and temporal sampling in order to obtain videos of uniform length. The training set consisted of 80% of the dataset, while the remaining 20% was for testing. Each model was trained with a batch size of 32 and an adaptive learning rate scheduler. The models were evaluated to ensure that we reliably and confidently measure standard metrics of performance for each of the deepfake types. The metrics included precision, recall, F1-score, and ROC-AUC. Tensorflow and Keras libraries were used for the model development, while OpenCV and FFmpeg were used to process the videos.[3] Alongside machine-learning experiments, the study involved a doctrinal legal analysis to explore the legal framework for synthetic media, misinformation and digital impersonation, in the context of developments in Indian and comparative international law. Key legal documents were, inter alia, the Indian Penal Code, 1860, the Information Technology Act, 2000, as well as several decisions by the Supreme Court of India in respect of privacy and digital identity rights.[4] The study also referenced other laws and their legislative history in the US and World, such as the DEEPFAKES Accountability Act, the AI Act proposed in the European Union, and what others have pointed to as regulatory gaps.[5] The purpose of the legal analysis, other than to understand the current context, was to clearly identify the restrictions and potential redress of current laws to be used in the case of harms relating to deepfakes.

Considering ethical issues was a core concern of this study. All datasets and the data itself were publicly available and freely and legally required for academic purposes. This study ensured that the aim of the experiments was not to misuse the person(s) involved or to misrepresent where the experiments lead. The methodology is focus on the aims of reproducibility and transparency, and some balance regarding the technological innovation of an AI model and the impact this could have in society.

Methodology Step-by-Step Procedure:

Preprocessing and Data Gathering:

The first operative step was to gather a diverse collection of sample videos-real and fake-from each database we've mentioned above. Then, everything was standardized with regard to frame rates, resolutions, and audio formats.

We focused on processing the area of the face through dlib's fifteen-point facial landmark detector. Further checks included multimodal alignment, synchronizing voice with lip movement.

Extracting features is thus very important in spotting deepfake signs, and the spectrogram analysis underlined some of them. Instances of pixel-level noise, patterns of blink rate, barely visible micro expressions in the face, and auditory inconsistencies, to be precise, appeared to be some of the indicators. We wanted to separate artificial signatures from genuine speech and facial movements.

For testing and training, we used deep learning architectures such as CNNs and Transformer-based approaches aiming to detect manipulations in the visual, audio, and temporal domains. The training phase was improved whenever we could by using data augmentation approaches that simulated environmental interference of low-resolution video, uneven lighting conditions, or camera shakes. In terms of validation, models were trained on one dataset and later tested on a second, enabling one to assess reliable robustness of such models. This cross-data accuracy, precision, r curves to evaluate. We also delved into the decision-making processes of the models by creating class activation visualizations and attention maps.

Finally, to ensure these detection models are ready for real-time forensic investigations, we integrated them into simulated environments like social media and law enforcement scenarios. We placed a strong emphasis on reproducibility, making sure that other researchers could achieve consistent results using the same datasets and codebase.

The Technology Behind Deepfakes

Deepfakes are perhaps the most discussed (and dreaded) product of artificial intelligence in the past few years. They employ advanced techniques in deep learning to generate highly realistic fake images, audio, or video of actual individuals. On first glance, deepfakes can appear to be just harmless amusement or clever tech wizardry, but they have come to pose a real concern to cybersecurity, digital trust, and public safety in a very short time.

It was in 2017 that the word "deepfake" first gained use when a Reddit user started posting these fake videos that replaced celebrity faces in sex content. The first wave left the world reeling — not for what was done, but for how convincingly the fakes were done. It became clear in a moment that anyone's identity could now be hacked with terrifying accuracy. Since then, deepfake technology has spread from entertainment or pranks into major threats in politics, the media, finance, and personal privacy. Experts now regard it as one of the most rapidly expanding dangers to cybersecurity and digital authenticity.

How It Works – The Technology Behind Deepfakes:

Central to deepfakes are sophisticated AI technologies, particularly Generative Adversarial Networks (GANs) and autoencoders.

Generative Adversarial Networks (GANs)

GANs function somewhat like a war between two minds:

The Generator attempts to generate bogus content (such as a video of someone uttering something they never actually uttered).

The Discriminator attempts to identify whether that content is genuine or bogus.

Both components continue training against one another, and the generator continues to get better until its content is so believable that even the discriminator can be tricked.

Autoencoders

Autoencoders are face translation machines. They:

Encode an individual's facial information into its compressed form.

Then transform that compressed form into a doctored face — usually transposed onto another person's body.

That is how deepfakes swapping faces are typically made.

How a Deepfake Video Is Made –

1. **Gathering Information**: You require numerous images or videos of the subject you wish to impersonate — from various angles, poses, and lighting sources.
2. **Preprocessing**: AI software then aligns faces, crops them, and normalizes the visuals for training.
3. **Training the Model**: Through GANs or autoencoders, the system learns how the subject's face changes, smiles, blinks, or speaks.
4. **Swapping the Face**: The trained model places the synthetic face onto a different person's body, frame by frame.
5. **Post-production Touches**: Color matching, blending edges, and smoothing out everything so it appears seamless and real.

Other Technologies Behind Deepfakes:

Facial landmark detection: Assists in monitoring facial landmarks in real time.

Lip-sync models such as **Wav2Lip**: Sync artificially created lips with actual speech seamlessly.

Voice cloning software like **Tacotron** or **WaveNet**: Resurrect someone's voice based on only a few minutes' worth of audio.

3D motion capture: Animates faces and bodies realistically, commonly applied in real-time deepfakes.

From Amusement to Terror – The Cybersecurity and Ethical Crisis:

Deepfakes were once considered a technological wonder, but now they are a time bomb. They've been employed to: Promote political propaganda by presenting leaders as saying things they never did. Perpetrate fraud by impersonating CEOs' voices or faces in financial crimes. Intrude on privacy, particularly through non-consensual explicit materials. Erode media trust, obscuring the distinction between authentic and fabricated news.

As deepfakes become more difficult to detect, the demand for counter-technologies increases. Applications such as Microsoft's Video Authenticator, Meta's Deepfake Detection Challenge, and AI-based forensic tools are filling in to push back. Methods such as watermarking, frequency analysis, and biometric monitoring are also surfacing to reveal digital forgeries.

But in this competition between creation and discovery, one thing is certain — the world will have to continue developing shields to safeguard truth in the era of AI.

Can AI Find DeepFake Videos?

AI can tell if something is a deepfake. That's why a lot of security experts agree that you need to use AI to fight AI. This is especially important now that deepfakes are getting better and it's harder to tell them apart from real ones. You can teach AI to find these kinds of changes and recognize voices or facial expressions that aren't normal. There are some ways to use AI to find deepfakes. One common way to do this is to use a two-step process: capture and analysis. In this process, you take pictures or videos of the real world and then check the content by looking at its most important details, like objects or facial features, to make sure they are real. In 2017, a Reddit moderator came up with the term "deepfake" and started a subreddit for it. This is when it became well-known in the news. At first, the idea was to make videos of famous people using face-swapping technology. However, it quickly grew into more complicated hoaxes, often with sexual content.

Results and Discussions

Results

In alignment with the aims stated in the abstract, the study successfully implemented and tested deepfake detection mechanisms using convolutional neural networks. XceptionNet emerged as the most effective model, achieving a detection accuracy of approximately 92.6%, while EfficientNet followed with a close 89.8%. These models were evaluated using established performance metrics, including precision, recall, F1-score, and cross-validation techniques, ensuring reliability and generalizability. The models were trained on datasets such as FaceForensics++, Celeb-DF v2, and the DFDC Preview Dataset, each offering real and synthetically altered video content.[1–3]

The extracted data was visualized using side-by-side comparisons of manipulated and authentic video frames. Notably, artifacts like lip-sync mismatches, irregular eye blinking, and skin tone inconsistencies under variable lighting were repeatedly identified by the models as strong indicators of manipulation.

Discussion

The results substantiate the critical concerns raised in the abstract regarding media authenticity and legal integrity. As this research demonstrates, the ability of AI-powered models to accurately flag deepfakes is promising—but not infallible. Even with accuracy nearing 90%, a 10% margin of error could have grave consequences in legal contexts, particularly when digital evidence is submitted in court.

Furthermore, the findings validate the need for stronger legal frameworks to regulate deepfake technology. The results highlight how manipulated media can bypass human detection and influence public opinion, criminal investigations, or even electoral outcomes—thus amplifying the urgency for ethical and legal oversight.

In addition, this section emphasizes the interdisciplinary importance of the research: not only are the technical achievements (like model performance) significant, but so is their legal relevance. The research supports the development of forensic tools that can be used by law enforcement agencies, legal practitioners, and digital forensic experts, contributing directly to the prevention and prosecution of deepfake-related offenses.

Ultimately, this study encourages future research to continue bridging the gap between technological innovation and legal preparedness, ensuring that as deepfakes evolve, society remains equipped to detect and deter them responsibly. Additionally, The results and discussion

What is the accuracy of deepfake detection models? The effectiveness of a deepfake detection model relies on the dataset and methods. In a controlled environment, AI-based models, such as CNN's, deep learning, and transformer architectures, have performance levels that range from 85% to 98%. Moral and Social Impact. This issue goes beyond technical issues with accuracy and precision of deepfake detection models, and has serious moral implications for current and future generations of society. Society is threatened by deepfakes being used in upstream information practices (i.e. propaganda) and non-consensual content. On the other hand, the use of deepfake technology has benefits in entertainment and improving accessibility (i.e. AI driven voice synthesis being used as a way to help recreate speeches of historic figures). We could be left wondering how to engage in a reflective and reality-assessing relationship with technology. Debates are still occurring about these positive and negative applications. Conclusion Deepfake technology is advancing rapidly, providing potential dangers and opportunities. The technological advances in detection have a lot of promise, but will need to keep pace with the rapid advances in AI. Future responsible applications of deepfake technology will be largely shaped by legal and ethical frameworks.

Conclusion

In summary, this research investigated the increasing threat of deepfakes and how Artificial Intelligence (AI) is inherently both the problem and the solution. As mentioned in my abstract, deepfake is a form of synthetic media created using advanced new AI technologies, particularly deep learning models, specifically Generative Adversarial Networks (GANs).[1] While the term deepfake was coined by a Reddit user in 2017,[2] it was first implemented in entertainment for face-swapping. However, deepfake technology quickly developed into a means for wrongful manipulation for malicious purposes, including creating non-consensual sexual content, along with political disinformation. The improvements to qualitatively producing deepfakes are such that one can often not even detect them. This has made it's deployment of AI-based detection methods not just useful, but absolutely necessary - as AI tools operate based on an ever-growing data set that can be trained to recognize distorted qualities such as strange facial expressions, misaligned lip movement, inconsistent blinking, or changed voice quality to classify suspicious altered content.[3] One of the most useful methods include two sequential approaches where a recorded event is first captured and then reviewed to identify discrepancies in certain key analogous visual and audiovisual features fossils of the original action.[4] These techniques in varying forms are now being used in multiple contexts, including national security, journalism, and social media, to protect the integrity of the digital capture.The implications of this are incredibly broad in scope and suggest a shift not only in technology but also in law and ethics when it comes to how we regulate our digital identities and the integrity of media. There are several ways forward, but first, it is important for governments and tech corporations to work together in the short term to establish regulatory policies, improve public education/awareness around media literacy, and provide funding for ethically-focused AI that can understand and explain its findings.[5] Deepfake detection should be viewed holistically; this is a cooperative responsibility. Alternatively, it may be more productive in the long run to implement sweeping reforms to limit access to these already-powerful technologies while also allowing the same

REFERENCES

References (Style based on Bluebook 21st Edition)

1. Hao Li, Justus Thies & Matthias Nießner, FaceForensics++: Learning to Detect Manipulated Facial Images, in Proceedings of the IEEE International Conference on Computer Vision (ICCV) 1 (2019).
2. Yuezun Li, Xin Yang, Pu Sun, Hongang Qi & Siwei Lyu, Celeb-DF: A Large-Scale Dataset for DeepFake Forensics, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 3207 (2020).
3. Andreas Rössler et al. FaceForensics: A Large-Scale Video Dataset for Forgery Detection in Human Faces, arXiv preprint arXiv:1803.09179 (2018) https://arxiv.org/abs/1803.09179
4. H.R. 3230, 116th Cong. (2019) (proposed DEEPFAKES Accountability Act).

5. European Commission, Proposal for a Regulation laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), COM (2021) 206 final (Apr. 21, 2021).

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

7. Seung Hyun Lee & Heesung Kwon, Detecting Deepfake Videos Using Convolutional Neural Networks, Electronic Imaging 2020, no. 5, (2020): 532-1.

8. Siwei Lyu, DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection, Cornell Tech Working Paper, https://arxiv.org/abs/2001.00179 (last visited June 15, 2025).

9. DeepFaceLab, GitHub Repository, https://github.com/iperov/DeepFaceLab (last visited June 15, 2025).

10. FaceSwap, GitHub Repository, https://github.com/deepfakes/faceswap (last visited June 15, 2025).

11. PyTorch Contributors, PyTorch: Deep Learning Framework, https://pytorch.org (last visited June 15, 2025).

12. OpenCV Library.

THE RIGHT TO ONE'S VOICE: COMBATTING AI-BASED IMITATION AND UNAUTHORIZED COMMERCIAL EXPLOITATION

Parimal Wagh & Saniya Ahir

Abstract

In the digital era, artificial intelligence has revolutionized content creation, enabling the replication and synthesis of human voices with unprecedented accuracy. While this technological advancement offers numerous benefits, it also raises significant legal and ethical concerns, particularly regarding the unauthorized use of an individual's voice for commercial gain. The right to one's voice is an essential aspect of personality rights, granting individuals control over how their identity is used, especially in commercial contexts. However, the advent of AI-powered voice cloning has led to rampant misuse, where voices of well-known personalities, including singers, actors, and public figures, are replicated without their consent for endorsements, advertisements, and entertainment. Such unauthorized usage not only infringes on publicity rights but also raises concerns about tarnishment and dilution, where AI-generated distortions harm the individual's reputation.

This paper examines the existing legal framework in India, particularly the Copyright Act, 1957, and its provisions under Section 38-B, which grants moral rights to performers. It also discusses the application of tort law, including the principles of misappropriation and dilution, to address AI-driven violations. Additionally, global precedents and evolving regulatory efforts are analyzed to provide a comparative perspective. Given the rapid advancements in AI, this paper argues for a more robust legal approach, including explicit statutory protections for voice rights, stringent enforcement mechanisms, and clear liability frameworks for AI developers *and users. Strengthening these safeguards is crucial to ensuring that individuals retain control over their voice and identity, preventing exploitation in the digital age. The research ultimately aims to contribute to the discourse on protecting human identity against the unchecked proliferation of AI-generated voice imitations.*

1. Introduction

As technology continues to advance globally, artificial intelligence ("**AI**") along with its foundational elements, machine learning and deep learning, has given rise to increasingly sophisticated applications. One notable innovation is speech synthesis, or Text-to-Speech ("TTS"), which has earned developing consideration in later a long time. This innovation changes over composed content into talked words and is broadly utilized in applications such as virtual colleagues and chatbots.

A major drawback of traditional speech synthesis is that its artificially generated voices often lack naturalness, sounding robotic or mechanical. However, the development of voice cloning has significantly improved this by allowing the creation of realistic speech that closely mimics a specific person's voice using a short audio sample. While this breakthrough has beneficial applications, it has also paved the way for the emergence of deepfake technology, which enables the manipulation of audio and video content. Previously, producing deepfake material required advanced technical expertise, but modern AI tools have made this capability more accessible to the general public. This increasing availability has sparked concerns over its potential misuse, making it essential to assess the current state of voice deepfake technology and examine both its advantages and risks.

Speech Synthesis or TTS is the computer-generated recreation of human discourse and alludes to the artificial change of content to sound. The objective of a great TTS framework is to have a computer do it, considering the instinctive nature and expressiveness of the voice. It could be a cutting-edge innovation within the field of information processing which includes numerous followers, such as, acoustics, linguistics, digital signal processing or computer science. A computer framework utilized for this reason is called a speech synthesizer and can be actualized in program or equipment. The quality of a speech synthesizer is judged by its similitude to the naturalness of a human voice and by its capacity to be caught on. Before electronic signal processing was designed, speech analysts attempted to construct mechanical machines to form human speech. In St. Petersburg 1779, the researcher Christian Kratzenstein, clarified the differences and built models of the five long vowels. This was taken after by von Kempelen of Vienna, in 1791, who included models of the tongue and lips, empowering the generation of consonants as well as vowels.

The very to begin with full electrical synthesis gadget was presented by Stewart in 1922. This machine was able to create single vowel sounds, but not any consonants or utterances. Within the 1930s, Bell Laboratories created the VOCODER, an electronic speech analyzer and synthesizer, at that point refined into the VODER. After this, the scientific world got to be more curious about speech synthesis since it was at long last appeared that intelligible speech may well be delivered artificially. In 1961, physicist John Larry Kelly used an IBM 704 computer to synthesize speech, which has gotten to be an event among the foremost noticeable within the history of Bell Labs. In fact, Kelly's voice synthesizer reproduced the melody “Daisy Bell”, which was used in the climactic scene of Arthur C. Clarke's screenplay 2001: A Space Journey.[1]

In late 1970's and early 1980's, a impressive number of commercial speech synthesis products were introduced and numerous computer operating systems have included speech synthesizers since then. The first speech system integrated into an operating system was Apple Computer's MacInTalk in 1984, displayed during the introduction of the Mac. The second operating system with advanced speech synthesis capabilities was AmigaOS, introduced in 1985, with both male and female voices. Speech systems were first available on Microsoft-based operating systems in Windows 95 and Windows 98.

As AI technologies, such as synthetic voice generation and deepfakes, become more prevalent, existing legal frameworks are being scrutinized for their adequacy in addressing the challenges posed by these advancements. This scrutiny has prompted governments and legislators worldwide to consider developing regulations specifically tailored to these emerging technologies. Furthermore, Section 57 of the Copyright Act of 19572 deals with authors special rights, known as "Moral rights". In the case of Genda Phool, a song remade from an old Bengali folk song sung by Ratan Khar was accused of disregarding the performance and moral rights of the singer.3

Prior to the 2012 amendment of the Copyright Act of 1957, singers in India did not possess specific rights over their performances and were not entitled to royalties. The amendment introduced "Performer's Rights" under Section 38 of the Act[4], granting performer's exclusive rights over their performances, including the right to receive royalties and the requirement for obtaining their consent for public use of their recordings. These rights are protected for 50 years following the year of the performance.[5]

2. Personality and Publicity Rights: Legal Foundations

Personality rights refer to an individual's authority to prevent the unauthorized use of their personal attributes, such as their name, image, voice, or likeness. This right includes both commercial and non-commercial uses, but its interpretation differs across legal systems. Some jurisdictions treat it as a single right covering both aspects, while others separate the commercial and non-commercial elements into distinct rights. As a result, terminology varies; in India, for example, "personality rights" and "publicity rights" are used interchangeably.[6] Personality rights refer to the legal entitlements individuals hold over their name, image, reputation, likeness, or other distinctive aspects of their identity, as well as any associated information. If an unauthorized third party attempts to derive commercial benefit from such attributes or information, it may constitute a violation of these rights.[7]

Personality rights are a critical component of legal systems worldwide, aimed at safeguarding an individual's control over their identity and personal attributes. These rights are particularly relevant in an era where digital platforms and media make it easier to misuse personal information, images, or likenesses without consent. While the core principle revolves around protecting individuals from unauthorized exploitation, the scope of these rights often extends to both economic and moral dimensions. Economically, personality rights prevent third parties from profiting from an individual's identity, such as through unauthorized endorsements or merchandise. Morally, they protect against harm to an individual's dignity, reputation, or privacy.

In certain instances, courts have interpreted the protection of personality rights in a manner analogous to the protection afforded to well-known trademarks under the Trade Marks Act, 1999 (India), particularly under Section 2(m)[8], which pertains to names and signatures. Additionally, the Copyright Act, 1957 provides relevant provisions for safeguarding personality rights. Section 2(qq) of the Copyright Act[9] defines a "performer" broadly to include actors, singers, musicians, dancers, and other individuals who deliver performances, potentially bringing personality rights within the scope of performer rights. Furthermore, Section 38 of the Copyright Act[10] grants performers the right to prevent unauthorized commercial exploitation of their performances. This provision effectively restricts the unauthorized marketing or use of an individual's performance, thereby offering a legal basis for protecting aspects of personality rights under copyright law. Copyright fosters creativity by guaranteeing creators the chance to secure financial benefits from their works.[11]

Another relevant factor is illustrated under Section 17(b) of the Copyright Act, 1957,[12] which, subject to the provisions of clause (a), regulates the unauthorized use of personality attributes falling within the scope of copyrighted works. This section establishes the rights of the first author or owner of the work. However, conflicts may arise between the individual who performed the work and the person who holds the rights over the performance, highlighting potential complexities in the application of this provision.[13]

3. AI and the Unauthorized Use of Voice

AI platforms are increasingly leveraging advanced algorithms to generate audio and visual content that replicates or mimics various aspects of an individual's identity such as a person's voice, singing style, photographs, images and other distinctive personality traits. One of the key technologies enabling this capability is Real Voice Cloning ("**RVC**"), which uses deep learning and neural networks to create highly accurate reproductions of a person's voice and other characteristics. While these advancements offer innovative opportunities in fields like entertainment, marketing, and education, they also raise significant legal and ethical concerns, particularly regarding the unauthorized use of an individual's identity.

The ability of AI to replicate personal attributes with precision poses challenges to existing legal frameworks, such as copyright law, personality rights, and privacy protections. For instance, the unauthorized use of a person's voice or likeness through RVC technology could infringe on their personality rights, which are designed to protect individuals from the exploitation of their identity without consent. Additionally, such practices may violate copyright laws if the replicated content is derived from protected works, such as recorded performances or artistic creations. The misuse of RVC and similar technologies could lead to identity theft, defamation, or the spread of misinformation, harming an individual's reputation and privacy. As AI continues to evolve, there is a pressing need for comprehensive legal frameworks and ethical guidelines to address these challenges, ensuring that technological advancements do not come at the expense of individual rights and societal trust.

In the case of *Karan Johar,*[14] the Hon'ble Bombay High Court ruled that personality rights, including the right to publicity, are inherently vested in celebrities. The judgment emphasized that the unauthorized use of a celebrity's name or other personal attributes constitutes a violation of their valuable personality rights and right to publicity.

In the landmark case of *R. Rajagopal v. State of Tamil Nadu*,[15] famously known as the Auto Shankar case, the Supreme Court of India recognized the right to privacy as a fundamental right under Article 21 of the Constitution, which guarantees the right to life and personal liberty. The court held that the right to privacy has two distinct but interconnected dimensions:

i) The right to privacy originated in tort law, providing individuals with a cause of action for damages resulting from the unlawful invasion of their privacy. This includes situations where a person's name, likeness, or life story is used without their consent, whether for advertising, non-advertising, or publishing purposes. Suchunauthorized use constitutes a violation of privacy, and the affected individual can seek legal remedies.

ii) The right to privacy also protects individuals from unlawful governmental intrusion into their personal lives. It encompasses the right to safeguard private matters such as family, marriage, procreation, motherhood, child-bearing, and education. No one can publish details about these aspects without the individual's consent, regardless of whether the content is truthful, laudatory, or critical. However, this protection does not apply if a person voluntarily thrusts themselves into a public controversy.

The court also outlined exceptions to this right. For instance, if information is based on public records, including court records, it can be published without violating privacy rights, as such matters are no longer considered private. However, the court carved out a significant exception to protect the dignity of women who are victims of sexual assault, kidnapping, abduction, or similar offenses. In such cases, the publication of the victim's name or the incident is prohibited, even if the details are part of public records, to prevent further indignity and harm.

4. Legal Frameworks Governing Voice Rights

The recognition of performers' rights under copyright law has undergone a gradual yet significant transformation. In the past, performers, whether actors, singers, dancers, or musicians were not granted distinct legal protections, as their contributions were often deemed unprotected labor. However, with the emergence of technologies such as sound recording and broadcasting, the necessity of safeguarding their creative efforts became evident. Over time, international agreements and amendments to national laws, including India's Copyright Act, progressively granted performers legal rights, ensuring they receive both financial compensation and moral recognition for their work.

Performers' rights were virtually unrecognized in India when the Copyright Act of 1957 was first introduced. It was only in 1978, with the *Fortune Films v. Dev Anand* case, that the issue came under judicial scrutiny. The Bombay High Court ruled that performers had no legal control over the use of their performances, as the Copyright Act did not expressly grant them such rights. This decision exposed a major loophole in Indian copyright law and sparked calls for legislative change.[16]

India’s recognition of performers’ rights took a significant step forward with the Copyright Amendment Act of 1994. This amendment introduced Section 38 into the Copyright Act, explicitly granting legal rights to performers. It also added Section 2(qq), which broadened the definition of a "performer" to include actors, singers, dancers, musicians, acrobats, jugglers, conjurers, snake charmers, lecturers, and others. This marked the first formal recognition of performers as rights holders under Indian copyright law.

Further developments came with the 2012 amendment, which introduced Sections 38-A and 38-B. These provisions granted performers both economic and moral rights. Economic rights allowed performers to control the recording, reproduction, and distribution of their performances, while moral rights ensured recognition as the performer and protection against any distortion or unauthorized modification of their work. In addition, performers were granted exclusive rights for a period of 50 years from the year of performance, providing them with long-term financial and legal protection. This amendment also empowered performers to claim royalties when their work was commercially exploited, ensuring they received fair compensation for their artistic contributions. Also, the amendments strengthened legal remedies for infringement, enabling performers to pursue civil and criminal actions. One such measure was the Anton Piller Order, which permitted performers to inspect and seize evidence in cases of copyright violations. These legal reforms significantly enhanced the protection of performers' rights in India [17]

Laws surrounding voice cloning and deepfake technology are still evolving to keep up with rapid advancements, but several legal frameworks already address related concerns. In the U.S., privacy laws such as the California Consumer Privacy Act and the Illinois Biometric Information Privacy Act offer some protection against the unauthorized use of personal data, including voice recordings. However, these laws vary by state and have certain limitations. Defamation laws also come into play when a cloned voice is used to spread false and damaging information. Depending on the format, it may be classified as either libel (written defamation) or slander (spoken defamation). Also, the right of publicity laws in states like California, New York, and Tennessee prevent the unauthorized commercial use of a person's identity, including their voice. These protections are particularly relevant for public figures and celebrities whose voices may be replicated without consent.[18]

The recognition of performers' rights has been a gradual but essential shift in copyright law, evolving alongside technological advancements that have transformed the entertainment industry. Initially overlooked, performers have fought for legal recognition, culminating in key amendments like those in India’s Copyright Act of 1994 and 2012, which granted them control over their work, financial benefits, and moral rights. However, the rise of digital manipulation, including voice cloning and deepfake technology, presents new challenges that existing laws struggle to fully address. While privacy and defamation laws offer partial protection, the rapid evolution of synthetic media demands stronger, more adaptive legal frameworks. Moving forward, safeguarding performers and individuals from unauthorized replication and exploitation will be crucial in preserving artistic integrity and personal identity in the digital era.

5. Judicial Precedents and Case Law Analysis

The evolution of personality rights has been significantly influenced by various judicial decisions across jurisdictions. These cases have addressed the unauthorized use of individuals' likenesses, voices, and personas, thereby shaping the legal landscape concerning personality rights.

Indian Cases

1. *Anil Kapoor vs. Simply Life India (2024)*

The Delhi High Court, in *Anil Kapoor's* case[19], affirmed that celebrities have significant rights over their name, image, likeness, and persona, which are crucial for their reputation and livelihood. The court held that unauthorized commercial use of these attributes, including on merchandise, ringtones, and GIFs, constitutes a violation of personality rights and trademark infringement. It also highlighted the growing misuse of AI to replicate a celebrity's voice and likeness, often in derogatory ways, such as on pornographic websites or through morphed content. In September 2023, Kapoor sought legal protection against AI-generated content exploiting his image and voice, including his iconic phrase "jhakaas." The court recognized personality rights as part of privacy, aligning with the Supreme Court's stance on privacy as a fundamental right. It issued an injunction preventing unauthorized use of Kapoor's persona, reinforcing legal protection against AI-driven exploitation and emphasizing the broader implications of technological misuse.

2. *Titan Industries Ltd. vs. M/s Ramkumar Jewellers (2012)*

In this landmark case, Titan Industries, the proprietor of the 'TANISHQ' brand, initiated legal action against Ramkumar Jewellers for the unauthorized use of advertisements featuring celebrities Amitabh Bachchan and Jaya Bachchan, who were official brand ambassadors for 'TANISHQ'. The defendant replicated these advertisements, leading to allegations of copyright infringement and misappropriation of personality rights. The Delhi High Court ruled in favor of Titan Industries, emphasizing that unauthorized commercial use of a celebrity's image without consent constitutes a violation of personality rights. This decision underscored the necessity of obtaining permission before utilizing an individual's persona in advertisements.[20]

International Cases

1. *Midler vs. Ford Motor Co. (1988, USA)*

In this case, singer Bette Midler sued Ford Motor Company for using a sound-alike to perform one of her songs in a commercial after she had declined to participate. The Ninth Circuit Court of Appeals held that intentionally imitating a celebrity's distinctive voice for commercial purposes without consent constitutes tortious misappropriation. The Court stated, "a voice is as distinctive and personal as a face," thereby extending legal protection to vocal likeness.[21] Television personality Vanna White sued Samsung for a commercial depicting a robot resembling her likeness in the context of the game show "Wheel of Fortune". The Ninth Circuit Court of Appeals ruled in favor of White, recognizing that the appropriation of a person's identity, even without using their exact likeness or name, can violate the right of publicity. The decision highlighted that the right of publicity protects against unauthorized commercial exploitation of one's persona.[22]

2. *Lohan vs. Take-Two Interactive Software Inc. (2018, USA)*

Actress Lindsay Lohan filed a lawsuit against Take-Two Interactive, alleging that a character in the video game "Grand Theft Auto V" was modeled after her without permission. The New York Court of Appeals dismissed the claim, stating that the game's character was a generic "twenty-something" woman without sufficient likeness to Lohan. The ruling clarified that incidental similarities do not constitute a violation of personality rights.[23]

The evolution of personality rights in India has been significantly influenced by landmark judicial decisions. In Titan Industries Ltd. v. M/s Ramkumar Jewellers (2012), the Delhi High Court addressed the unauthorized use of celebrities Amitabh and Jaya Bachchan's images in advertisements by the defendant, ruling such actions as violations of personality rights and emphasizing the necessity of obtaining consent before utilizing an individual's persona for commercial purposes. More recently, in Anil Kapoor v. Simply Life India & Ors. (2023), the Delhi High Court granted an interim injunction protecting Bollywood actor Anil Kapoor's personality rights against unauthorized AI-generated content that misused his likeness, voice, and the iconic phrase "jhakaas." The court's decision underscored the importance of safeguarding individuals' persona in the digital age, setting a precedent for protecting celebrity rights against AI misuse. Internationally, cases like Midler v. Ford Motor Co. (1988) and White v. Samsung Electronics America, Inc. (1992) have reinforced the protection of personality rights.

Collectively, these cases highlight a judicial trend towards safeguarding individuals' personality rights, emphasizing the necessity of obtaining consent before utilizing an individual's persona for commercial gain. This progression reflects a broader commitment to protecting personal identity in an era where technological advancements make unauthorized exploitation increasingly feasible.

6. Challenges in Regulating AI-Generated Voice Imitation

Voice deepfake content raises significant concerns as it infringes on copyright, particularly affecting singers and voice-over artists. These individuals often face challenges in legally enforcing their rights, as copyright for a song is granted as a whole, with ownership typically vested in the producers.[24] Voice cloning poses a serious threat, as it can strip individuals of their identity, enabling misuse for fraud, defamation, or reputational harm. It is crucial to distinguish between synthetic media and deepfakes, as synthetic media encompasses broader applications of this technology that can be used positively. For instance, in the entertainment industry or in medical contexts, synthetic media can help individuals who have lost their voices due to surgeries or other conditions.[25]

Recently, a Berkeley-based AI start-up allegedly stole a voice actor's voice to train its AI software under the pretence of using it solely for research purposes. Such practices not only violate intellectual property rights but also undermine trust in the ethical use of emerging technologies.[26] Significant ethical concerns accompany the advancements in AI, particularly in relation to identity theft, defamation, and the misuse of synthesized content. The potential for harm is substantial, as AI-generated content can be used to manipulate or misrepresent individuals, causing reputational damage or financial loss. To mitigate these risks, it is imperative to establish clear ethical guidelines and industry practices that prioritize consent, transparency, and accountability in the use of AI technologies.[27] In a notable case[28] involving the personality rights of a deceased individual, the Delhi High Court ruled that personality rights are not heritable and expire upon the death of the individual. This decision underscores the limitations of current legal frameworks in addressing posthumous rights and highlights the need for legislative clarity on the protection of personality rights in the context of AI and digital technologies.

Beyond commercial and creative harm, voice deepfakes also raise serious privacy and reputational concerns. The ability to replicate someone's voice with precision using AI poses a threat to personal identity and security.

Misuse of this technology can lead to fraud, defamation, and other forms of exploitation, emphasizing the urgent need for robust legal and ethical safeguards to protect individuals from such violations.[29]

7. The Need for Legislative Reforms

In India, performers such as artists are granted rights under Sections 38A and 38B of the Copyright Act of 1957, which protect their visual or acoustic performances. However, these provisions are limited to individuals who actively engage in performance, as defined under the Act. These provisions fall short in addressing situations where voices are replicated using AI, as AI-generated voice cloning does not involve an actual performance by the individual. Since the Act does not extend performer's rights to such scenarios, individuals whose voices are cloned through AI must rely on personality rights for protection.[30] Personality rights safeguard an individual's unique identity attributes, such as their voice, from unauthorized use, making it the only viable legal recourse in cases of AI-generated voice replication.

India is in the process of drafting the Digital India Bill, which aims to comprehensively address the regulation of AI. The existing Information Technology Act, 2000 does not define AI or regulate its practices and processes, leaving a significant gap in the legal framework. Generative AI tools, while innovative, pose risks to human rights, national sovereignty, and integrity, necessitating stringent oversight. The IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, though a step forward, primarily focus on due diligence for intermediaries and lack a proactive approach to emerging technologies. The principle of 'safe harbour', which protects intermediaries from liability, requires reevaluation to ensure greater accountability in the digital space. As AI continues to evolve, India must adopt forward-looking legislation to address the challenges of the

21st-century digital revolution.[31] A robust legal framework is essential to balance innovation with accountability, safeguarding both individual and national interests in an increasingly digital world. The Digital India Bill is expected to fill these gaps, providing a comprehensive approach to AI regulation and addressing the complexities of modern technology.

8. Conclusion

The rapid advancement of AI has enabled the replication of human voices with striking accuracy, raising serious concerns about unauthorized commercial use and ethical implications. While India's Copyright Act, 1957, extends certain protections through performers' rights and moral rights under Sections 38-A and 38-B, these provisions are inadequate in dealing with AI-generated voice cloning. The fundamental limitation lies in the fact that these laws protect actual performances, whereas AI-generated voices do not require the individual's direct involvement. Consequently, individuals find themselves vulnerable to exploitation, including unauthorized commercial endorsements, defamatory use, and reputational harm.

Existing legal frameworks, such as personality and publicity rights, provide some level of protection but lack uniform interpretation across jurisdictions. Courts have recognized the need to safeguard an individual's persona through landmark rulings, such as the recent Anil Kapoor case in India and international precedents like Midler v. Ford Motor Co. However, while these cases emphasize the importance of protecting an individual's voice from misuse, they do not comprehensively address the challenges posed by AI-driven voice cloning, which continues to proliferate at an alarming rate. The absence of statutory protection specifically for voice rights further complicates matters, leaving victims with limited legal recourse.

The lack of a clear liability framework for AI developers and users exacerbates the problem. Without strict legal obligations on those who create and deploy AI-generated voices, accountability remains ambiguous, making it difficult for affected individuals to seek justice. While synthetic voice technology has legitimate applications, such as aiding people who have lost their ability to speak, the potential for abuse remains a significant concern. When misused, AI-generated voices can erode personal trust, mislead the public, and even contribute to identity fraud. To mitigate these risks, India needs a robust legal framework explicitly addressing AI-generated voice cloning. Legislative measures should define voice rights clearly, establish consent-based mechanisms for voice replication, and impose strict liability on those who engage in unauthorized use. The introduction of technological safeguards, such as watermarking AI-generated voices, could help distinguish between real and synthetic speech. As AI continues to evolve, ensuring the ethical and lawful use of synthetic voices is crucial to balancing innovation with individual rights and privacy.

References

1. AI voice cloning through the lens of copyright laws: challenges on the rights of singers https://www.theippress.com/2024/06/09/ai-voice-cloning-through-the-lens-of-copyright-laws-challenge s-on-the-rights-of-singers/
2. Augustian, A. (2023). Protection of personality rights in India: Issues and challenges. IPR Journal of Maharashtra National Law University, Nagpur, 1(1), 44–53. Retrieved from https://www.nlunagpur.ac.in/PDF/Publications/5-Current-Issue/4.%20PROTECTION%20OF%20PER SONALITY%20RIGHTS%20IN%20INDIA.pdf
3. Kaur, B., & Chauhan, G. (n.d.). Privacy and publicity: The two facets of personality rights. Remfry & Sagar. Retrieved from https://www.remfry.com/wp-content/uploads/2017/11/privacy-and-

publicity-the-two-facets-of-personality-rights.pdf

4. Trade Marks Act, section 2(m) - Acts of Parliament- 1999(India) -name and signature, Copyright Act, section 2 (qq) Acts of Parliament- 1957 (India).
5. Copyright Act, 1957 (Act No. 14 of 1957), s. 2(qq), Government of India.
6. Copyright Act, 1957 (Act No. 14 of 1957), s. 38, Government of India.
7. Apfelbaum, M. J. (1983). Copyright and the right of publicity: One pea in two pods. Georgetown Law Journal, 71, 1567.
8. Copyright Act, 1957 (Act No. 14 of 1957), s. 17(b), Government of India.
9. Augustian, A. (2023). Protection of personality rights in India: Issues and challenges. IPR Journal of Maharashtra National Law University, Nagpur, 1(1), 44–53. Retrieved from https://www.nlunagpur.ac.in/PDF/Publications/5-Current-Issue/4.%20PROTECTION%20OF%20PER SONALITY%20RIGHTS%20IN%20INDIA.pdf
10. Karan Johar (also known as Rahul Kumar Johar) v. Indian Pride Advisory Pvt. Ltd. & Ors., Interim Application (L) No. 17865 of 2024 in Commercial IPR Suit (L) No. 17863 of 2024.
11. R. Rajagopal v. State of Tamil Nadu (1994) 6 SCC 632
12. White vs. Samsung Electronics America, Inc. (1992, USA)
13. Titan Industries Ltd. vs M/S Ramkumar Jewellers, 2012 SCC OnLine Del 2382
14. Midler v. Ford Motor Co., 849 F.2d 460 (9th Cir. 1988)
15. https://law.justia.com/cases/federal/appellate-courts/F2/849/460/37485/
16. Fortune Films International vs Dev Anand And Anr., 1978 SCC OnLine Bom 156
17. What is the significance of performer's rights in the copyright regime?
18. https://legal-wires.com/lex-o-pedia/what-is-the-significance-of-performers-rights-in-the-copyright-regim
19. Understanding Voice Cloning: The Laws and your Rights. Retrieved

from https://www.nationalsecuritylawfirm.com/understanding-voice-cloning-the-laws-and-your-rights/ 19 Anil Kapoor v. Simply Life India, 2023 SCC OnLine Del 6914

20. White vs. Samsung Electronics America, Inc.(1992, USA), 26 U.S.P.Q.2D (BNA) 1362. https://www.casebriefs.com/blog/law/torts/torts-keyed-to-epstein/privacy/white-v-samsung-electronics- america-inc/
21. Lohan vs. Take-Two Interactive Software Inc. (2018, USA)
22. https://law.justia.com/cases/new-york/court-of-appeals/2018/24.html
23. Mondaq. (n.d.). Navigating voice deep fakes: Legal challenges framework. https://www.mondaq.com/india/new technology/1497122/navigating-voice-deep-fakes-legal-challenges-and-framework
24. Respeecher. (n.d.). Ethics in AI: Making voice cloning safe. https://www.respeecher.com/news/ethics-in-ai-making-voice cloning-safe#:~:text=However%2C%20significant%20ethical%20concerns%20exist,and%20practices%20within%20the%20industry.
25. CNN. (2024, May 17). Voice actors sue AI company Lovo for allegedly stealing their voices. https://edition.cnn.com/2024/05/17/tech/voice-actors-ai-lawsuit-lovo/index.html
26. Respeecher. (n.d.). Ethics in AI: Making voice cloning safe. https://www.respeecher.com/news/ethics-in-ai-making-voice-cloning-safe#:~:text=However%2C%20significant%20ethical%20concerns%20exist,and%20practices%20within%20the%20industry.
27. Krishna Kishore Singh v. Sarla A Saraogi, 2023 SCC OnLine Del 3997
28. Mondaq. (n.d.). Navigating voice deep fakes: Legal challenges and framework. https://www.mondaq.com/india/new-technology/1497122/navigating-voice-deep- fakes-legal-challenges-and-

framework

29. SpicyIP. (2024, September). Voice clones and legal tones: The intersection of artificial intelligence and posthumous personality rights. Retrieved from https://spicyip.com/2024/09/voice clones-and-legal-tones-the-intersection-of-artificial-intelligence-and- post-humous-personality-rights.html

30. National Strategy Institute. (n.d.). Why India urgently needs a legal framework to regulate artificial intelligence. Retrieved from https://www.natstrat.org/articledetail/publications/why-india-urgently-needs-a-legal-framework-to-regul ate-artificial-intelligence-173.html

THE ROLE OF AI IN JUDICIAL PROCESSES: ENHANCING EFFICIENCY, ACCURACY, AND ACCESS TO JUSTICE

Kaushiki Srivastava and Rishabh Srivastava[1]

Introduction

The emergence of Artificial Intelligence (AI) as a transformative force in governance and legal frameworks represents one of the landmark technological developments of the 21st century. AI systems, defined broadly as computational models that can perform tasks typically requiring human intelligence, have evolved from theoretical constructs to practical applications that are reshaping fundamental institutions of society, including the judiciary. This technological evolution has occurred against a backdrop of increasing digitisation across all sectors, creating both opportunities and challenges for traditional legal systems worldwide.

The increasing interest in utilising AI to rectify inefficiencies and injustices within justice systems stems from several converging factors. First, judicial systems globally face significant backlogs and delays, with cases sometimes taking years or even decades to resolve. In India, for example, approximately 47.2 million cases remain pending across various courts, with an estimated 41% of cases in high courts pending for five years or more.[2]

Similarly, in many jurisdictions, the cost of litigation has become prohibitively expensive for average citizens, creating a substantial justice gap between those who can afford legal representation and those who cannot.

[1] LL. b 2nd Year Students, New Law College, Bharati Vidyapeeth

[2] Sushmita Khandbahale, Can Technology Defuse the Ticking Pendency Bomb for the Indian Judiciary System? VEDHAS L.J., June 2024, at 45-67.

Second, inconsistencies in judicial outcomes where similar cases produce divergent results based on factors such as geographic location, individual judicial preferences, or implicit biases have undermined public confidence in the impartiality and reliability of legal systems. Research has repeatedly demonstrated the influence of extra-legal factors on judicial decision-making, including the timing of decisions (such as pre- or post-lunch rulings), demographic characteristics of defendants, and even weather conditions.[3]
Against this backdrop, AI technologies offer potential solutions to these persistent challenges. Machine learning algorithms can process and analyse vast datasets of legal information at speeds and scales unattainable by human legal professionals. Natural language processing can extract relevant information from unstructured legal texts, allowing for more comprehensive legal research. Predictive analytics can identify patterns in legal outcomes, potentially enhancing consistency and fairness. Automated administrative systems can streamline procedural aspects of justice delivery, reducing delays and backlogs.

Scope of Study

This study concentrates specifically on the judicial process rather than administrative or private legal technology applications. While the broader legal technology ecosystem encompasses a wide range of applications from contract management systems used by corporate legal departments to compliance monitoring tools employed by regulatory agencies, our focus remains on technologies that directly impact how courts and judicial officers execute their core functions of adjudication, case management, and justice delivery.

[3] Konstantina Terzidou, The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial, 18 UTRECHT L. REV. 121 (2022).

The judicial process involves various stages, from case filing to judgment enforcement. AI plays a growing role across this spectrum, particularly in enhancing procedural efficiency, improving decision accuracy, and promoting equitable access to justice. By automating administrative tasks, supporting legal research, and managing cases, AI helps reduce delays and backlogs. It also aids in evidence evaluation, legal analysis, and consistency checks to improve the quality and fairness of decisions. Furthermore, AI tools can lower legal costs, simplify procedures, and bridge linguistic or geographic gaps, ensuring broader access to justice, especially for underserved or marginalised communities.

By concentrating on these three dimensions, the study aims to provide a comprehensive assessment of AI's potential to transform judicial systems while acknowledging the technological, ethical, and institutional constraints that may limit or shape this transformation.

Significance of The Study

The potential for AI to revolutionise justice delivery in overwhelmed systems, such as in India, cannot be overstated. In countries with significant judicial backlogs, AI offers promising avenues for addressing systemic inefficiencies that undermine the fundamental principle *that justice delayed is justice denied*.

The Supreme Court of India has already initiated steps in this direction with the implementation of **SUPACE (Supreme Court Portal for Assistance in Court Efficiency),** an AI tool designed to assist judges in reading, summarising, and researching case materials.[4] Such innovations hold

[4] Shalmoli Basu & Chaitra Jha, Evaluating ICT Adoption in the Indian Judiciary: Challenges, Opportunities, and the Impact of the E-Courts Project, 15 INDIAN J.L. & JUST. 78 (2024).

particular significance in jurisdictions where judicial resources are stretched thin relative to caseloads.

The study is also significant in its examination of how AI can assist judges, lawyers, and litigants by providing timely, data-driven support across various stages of the judicial process. For judges, AI tools can offer research assistance, document summarisation, precedent identification, and consistency checking, potentially enhancing the quality and efficiency of judicial decision-making. For lawyers, AI-powered legal research platforms, document drafting assistance, and predictive analytics can augment professional capabilities and reduce routine workloads. For litigants, especially those without professional legal representation, AI chatbots, self-help tools, and simplified interfaces can make legal systems more navigable and responsive.

Finally, the study contributes to an emergent body of scholarship at the intersection of law, technology, and governance. Examining concrete applications, empirical outcomes, and normative implications of AI in judicial contexts helps bridge the gap between theoretical discussions about AI governance and practical implementations in real-world institutional settings. This applied focus is particularly valuable given the rapid pace of technological change and the need for evidence-based approaches to institutional innovation.

Enhancing Efficiency In Judicial Procedures

1. AUTOMATION OF ADMINISTRATIVE & PROCEDURAL TASKS

The integration of AI into judicial systems begins with the automation of administrative and procedural tasks that, while essential to judicial functioning, consume significant resources without directly contributing to substantive legal analysis or decision-making.

AI-enhanced case management systems go beyond simple digitisation by incorporating intelligent features that optimise judicial workflows. For example, machine learning algorithms can analyse case characteristics to predict resource requirements, estimate time to resolution, and prioritise cases based on urgency or complexity. In India, the e-Courts project has implemented electronic case management across various levels of the judiciary, with approximately 18,735 courts computerised as of 2023.[5] The system allows for online case filing, status tracking, and document submission, significantly reducing administrative burdens and physical paperwork.

The implementation of chatbots for procedural guidance represents another frontier in judicial efficiency enhancement. Court-operated chatbots can provide 24/7 assistance to litigants and legal professionals, answering frequently asked questions about court procedures, document requirements, filing deadlines, and fee structures. The United States Administrative Office of the U.S. Courts has developed **CLARA (Court Legal Assistance and Resource Agent)**, a chatbot that guides users through federal court procedures and forms. Similarly, the New South Wales Local Court in Australia has implemented a chatbot that assists self-represented litigants in navigating small claims procedures.

- **Case Study: SUPACE (Supreme Court Portal for Assistance in Court Efficiency), India**

The Supreme Court Portal for Assistance in Court Efficiency (SUPACE) represents India's pioneering effort to integrate AI into its apex judicial

[5] Shalmoli Basu & Chaitra Jha, Evaluating ICT Adoption in the Indian Judiciary: Challenges, Opportunities, and the Impact of the E-Courts Project, 15 INDIAN J.L. & JUST. 78 (2024).

institution. Launched in April 2021, SUPACE was conceptualised as a response to the overwhelming case backlog and research burden faced by Supreme Court judges and their legal research teams. The system is developed by the E-committee of the Supreme Court in collaboration with technical partners, including academic institutions and technology companies. The AI-enabled assistive tool can augment the efficiency of legal researchers and judges working on cases by extracting relevant information from case documents, identifying key legal principles and precedents, and generating preliminary draft summaries for judicial review.[6]

The system's functionality encompasses several key capabilities:

- **Document Analysis and Summarisation:** SUPACE can process lengthy legal documents, including petitions, counter-affidavits, and judgments, to extract key facts, legal arguments, and reliefs sought. The system uses natural language processing to identify the central issues in a case and generate concise summaries, potentially reducing the time judges spend on document review.
- **Legal Research Assistance:** The system can search through vast databases of case law, statutes, and scholarly articles to identify relevant legal precedents and principles applicable to a given case. By automating preliminary legal research, SUPACE aims to enhance the comprehensiveness of legal analysis while reducing research time.
- **Case Categorisation and Prioritisation:** SUPACE incorporates algorithms that analyse case characteristics to categorise matters based on subject matter, complexity, urgency, and potential impact. This functionality supports more efficient case allocation and scheduling.

SUPACE incorporates algorithms that analyse case characteristics to categorise matters based on subject matter, complexity, urgency, and potential

[6] Shami Sumer Singh, Use of AI-Driven Support System to Help Courts in Predicting Child Custody Outcomes in India, 18 WORLD J. ADVANCED RSCH. & REVS. 2108 (2024).

impact. This functionality supports more efficient case allocation and scheduling.

2. Legal Research And Drafting With Ai

AI-powered legal research platforms represent one of the most mature and widely adopted applications of artificial intelligence in the legal domain. These platforms leverage natural language processing, machine learning, and information retrieval techniques to transform how legal professionals identify, analyse, and apply relevant legal materials.

ROSS Intelligence, often described as ***"the world's first digital legal expert,"*** utilised IBM's Watson technology to process natural language queries and search through vast legal databases to identify relevant precedents, statutes, and scholarly articles.

CaseMine, developed in India but now used internationally, employs machine learning algorithms to identify conceptual similarities between cases beyond mere keyword matching. The platform's ***"CaseIQ"*** feature allows users to upload case documents or briefs and automatically identifies relevant precedents based on factual and legal similarities. *CaseMine* also offers visual representations of legal doctrine evolution and precedential relationships, enabling lawyers to better understand how legal principles have developed over time.[7] The impact of these AI-powered research platforms on judicial efficiency is multifaceted. First, they significantly reduce research time, with some studies suggesting time savings of 30-70% compared to traditional research methods.[8]

[7] Shalmoli Basu & Chaitra Jha, Evaluating ICT Adoption in the Indian Judiciary: Challenges, Opportunities, and the Impact of the E-Courts Project, 15 INDIAN J.L. & JUST. 78 (2024).

[8] S. Rohana, Transforming Legal Practice: The Rise of AI for Efficiency and Access to Justice, 4 INT'L J. LEGAL RSCH. 87 (2024).

Several courts have begun experimenting with AI-assisted drafting tools. The Administrative Office of the United States Courts has developed systems that generate draft orders for routine matters such as scheduling, continuances, and uncontested motions. Similarly, the Federal Court of Australia has implemented an ***"Auto-Text"*** system that helps judges prepare standardised components of judgments, such as case summaries and procedural histories.

The integration of these research and drafting technologies into judicial workflows represents a significant opportunity to enhance efficiency without compromising quality. By automating information retrieval and document generation tasks, AI tools enable legal professionals to allocate more time and cognitive resources to tasks that genuinely require human judgment, potentially leading to better-reasoned and more timely judicial outcomes.

3. AI Assisted Adjudication

Estonia, renowned for its digital governance initiatives, launched a pioneering AI Judge pilot project aimed at resolving small claims disputes efficiently. This initiative emerged from Estonia's broader e-governance framework, which includes the X-Road data exchange platform, e-residency program, and digital signature infrastructure. The AI Judge project specifically targeted small claims disputes under €7,000, an area where case volumes are high but legal complexity is typically limited.

It is important to note that despite significant media coverage suggesting otherwise, the Estonian Ministry of Justice has officially clarified that ***"Estonia does not develop an AI robot judge for small claims procedure nor general court procedures to replace the human judge"*** (Estonia Ministry of Justice and Digital Affairs, 2019).[9] While the project initially explored the

[9] Ministry of Just. & Digital Affs., Republic of Est., Estonia Does Not Develop AI Judge (Mar. 26, 2019), https://www.justdigi.ee/en/news/estonia-does-not-develop-ai-judge

potential for AI to adjudicate certain cases, it has evolved into a decision-support system rather than an autonomous judicial entity.

The actual implementation involves a semi-automated process for small claims and maintenance decisions whereby computer-generated payment orders are produced based on standardised inputs. The system analyses case information submitted through structured online forms, applies relevant legal rules, and generates preliminary decisions for judicial review. These preliminary decisions include assessments of jurisdiction, applicable law, basic legal reasoning, and proposed judgments.[10]

Estonia's approach represents a measured integration of AI into judicial processes, where technology augments rather than replaces human judgment. By carefully delineating the boundaries of automation and preserving human review options, the Estonian model suggests how judicial efficiency can be enhanced while maintaining essential procedural protections.

4. Improving Accuracy In Legal Outcomes

<u>4.1. USE OF THE PREDICTIVE ANALYTICS AND RISK ASSESSMENT TOOLS</u>

Predictive analytics and risk assessment tools represent one of the most controversial yet potentially transformative applications of AI in judicial systems. These technologies leverage machine learning algorithms to analyse historical data and identify patterns that can predict future events or outcomes. In judicial contexts, such tools have been primarily applied to assess recidivism rise the likelihood that an individual will reoffend after release, to inform bail, sentencing, and parole decisions.

[10] Kärt Pormeister, AI Systems' Impact on the Recognition of Foreign Judgements, 32 JURIDICA INT'L 107, 112 (2023).

In the United States, tools like ***COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)*** have been widely adopted across jurisdictions to assist judges in making pretrial detention and sentencing determinations. COMPAS analyses various factors, including criminal history, social demographics, and survey responses, to classify defendants into risk categories based on their predicted likelihood of reoffending or failing to appear for court dates. Similar tools include the ***Public Safety Assessment (PSA)*** developed by the Laura and John Arnold Foundation and the ***Virginia Risk Assessment Instrument (VRAI).***

Proponents argue that these tools can enhance decision quality by providing empirically grounded risk assessments that may complement or correct for potential biases in human judgment. Moreover, questions about algorithmic transparency and explainability have emerged, as many risk assessment tools employ complex machine learning models whose decision-making processes are not easily interpretable by human users. This "black box" nature raises procedural justice concerns, as defendants may be unable to effectively challenge assessments, they believe are inaccurate or unfair

- **Case Study: COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), USA**

COMPAS, developed by Northpointe (now Equivant), represents one of the most widely used and extensively studied risk assessment tools in the American criminal justice system. Implemented across numerous states, including Wisconsin, California, Florida, and New York, COMPAS assesses both general and violent recidivism risk by analysing responses to a 137-item questionnaire along with criminal history data. The system classifies defendants into low, medium, or high-risk categories, which judges may consider when making bail determinations, sentencing decisions, and parole evaluations.

However, COMPAS has faced significant criticism regarding racial bias and transparency issues. In 2016, ProPublica published an influential investigation that analysed COMPAS assessments of over 7,000 defendants in Broward County, Florida. The analysis found significant racial disparities in false positive rates: among defendants who did not reoffend within two years, Black defendants were nearly twice as likely to be misclassified as high-risk (44.9%) compared to white defendants (23.5%).[11] This disparity raised serious concerns about the tool's fairness and potential to exacerbate existing racial inequalities in criminal justice outcomes.

The ProPublica findings sparked an intense scholarly debate about how algorithmic fairness should be defined and measured. Northpointe contested ProPublica's analysis, arguing that COMPAS satisfied a different fairness criterion: similar risk scores represented similar recidivism probabilities across racial groups.[12] Subsequent research has demonstrated that different mathematical definitions of fairness (such as calibration, equal false positive rates, and equal false negative rates) cannot be simultaneously satisfied for groups with different base rates of the predicted outcome, in this case, recidivism rates.

[11] Timothy O'Brien, Compounding Injustice: The Cascading Effect of Algorithmic Bias in Risk Assessments, 13 GEO. J.L. & MOD. CRITICAL RACE PERSPS. 39 (2021).

[12] Anthony W. Flores et al., False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks," 80 FED. PROB. 38 (2016).

The landmark case of ***Loomis v. Wisconsin***[13] Brought these transparency issues to the forefront. Eric Loomis challenged his sentence, which was partly based on a COMPAS assessment, arguing that the use of a proprietary algorithm violated his due process rights. While the Wisconsin Supreme Court ultimately upheld the use of COMPAS, it imposed significant limitations, requiring that risk assessments be accompanied by written warnings about their limitations and prohibiting risk scores from being the determinative factor in sentencing decisions

.

5.2. Support In Judgment Consistency And Precedent Mapping

AI systems can significantly enhance judicial decision-making by identifying inconsistencies in legal outcomes across similar factual scenarios and assisting in the compilation of databases of consistent precedents for lower courts.

Precedent mapping tools employ natural language processing and machine learning techniques to analyse large corpora of judicial decisions, identifying patterns, similarities, and potential conflicts. These systems can detect conceptual relationships between cases beyond simple keyword matching, recognising when different courts have reached divergent conclusions on legally analogous facts. For example, the Indian startup CaseMine developed ***"CaseRanker,"*** which analyses judicial opinions to identify clusters of factually and legally similar cases and highlights inconsistencies in outcomes or reasoning.[14]

While these technologies offer significant potential benefits for judicial consistency, they also raise important questions about the nature of precedent and judicial discretion. Legal rules are not simply applied mechanically to

[13] 881 N.W.2d 749 (Wis. 2016)

[14] Shalmoli Basu & Chaitra Jha, Evaluating ICT Adoption in the Indian Judiciary: Challenges, Opportunities, and the Impact of the E-Courts Project, 15 INDIAN J.L. & JUST. 78 (2024).

facts but require interpretation and judgment. The appropriate balance between consistency and context-sensitivity in judicial decision-making remains a fundamental question that technology alone cannot resolve.

5.3. Challenge To Accuracy

Despite their potential benefits, AI systems in judicial contexts face significant challenges related to algorithmic opacity, explainability, and the impact of biased datasets on output quality.

Algorithmic opacity refers to the "black box" nature of many advanced AI systems, particularly those based on deep learning or complex machine learning models. These systems often develop internal representations and decision processes that are not readily interpretable by human observers, including their developers. In judicial contexts, where decisions must be justified and subject to review, this opacity creates fundamental tensions with core legal principles. When judges rely on algorithmic recommendations or analytics but cannot fully explain how those outputs were generated, the legitimacy of resulting decisions may be compromised.

The impact of biased datasets on output quality represents perhaps the most profound challenge to accuracy in AI-augmented judicial processes. Machine learning systems learn from historical data, and when that data reflects past biases or discriminatory practices, algorithms may encode and perpetuate these patterns. This is particularly concerning in criminal justice applications, where historical data may reflect systemic biases related to race, socioeconomic status, and other protected characteristics.

For example, if an algorithm is trained on historical sentencing data that reflects disparate treatment of minority defendants, it may reproduce these disparities in its recommendations, even without explicitly considering race as an input variable. Similarly, if an algorithm learns from patterns of police activity that disproportionately target certain neighbourhoods or communities, it may generate higher risk assessments for individuals from those areas, irrespective of their actual conduct.

Additionally, procedural safeguards can mitigate accuracy risks through careful specification of how AI tools should be used in judicial processes. In ***Loomis v. Wisconsin***[15] The court-imposed requirements that risk assessment scores be accompanied by written warnings about their limitations and prohibited such scores from determining sentencing outcomes on their own. Similar approaches, which position AI as decision support rather than decision replacement, can preserve the benefits of algorithmic analysis while maintaining essential human judgment and oversight.

5. Enhancing Access To Justice For Underserved Populations

5.1. _Reduction In Legal Costs And Barriers

Access to justice remains a pervasive challenge globally, with legal services often financially inaccessible to significant portions of the population. ***"DoNotPay",*** launched in 2015 as a chatbot to contest parking tickets, has evolved into what it terms an "AI Consumer Champion" with capabilities across multiple legal domains. The platform enables users to generate legal documents, write demand letters, cancel subscriptions, and navigate small claims procedures through a user-friendly interface that requires no specialised legal knowledge. By removing the need for attorney involvement in straightforward legal matters, DoNotPay has significantly reduced costs for users, advertising a $36 annual subscription compared to hourly attorney rates that often exceed $300 (DoNotPay, 2024).[16]

[15] 881 N.W.2d 749 (Wis. 2016)
[16] DONOTPAY, https://donotpay.com (last visited Apr. 14, 2025).

However, it's important to note that DoNotPay has faced regulatory scrutiny over some of its claims. In February 2025, the U.S. Federal Trade Commission finalised an order requiring DoNotPay to "stop making deceptive claims about the abilities of its AI chatbot" and pay $193,000 in monetary relief. The FTC action highlighted the tension between expanding access to legal services and ensuring that consumers receive accurate information about technological capabilities (Federal Trade Commission, 2025).[17]

More robustly, a study of JusticeBot's implementation in Quebec found that users who received guidance through the platform were significantly more likely to follow through with filing claims compared to those who only accessed static information resources.[18]

While these tools offer substantial promise for expanding access, they also face important limitations. Their effectiveness typically depends on internet access and digital literacy, potentially excluding the most vulnerable populations. Additionally, they work best for standardised legal issues with clear procedural pathways, whereas complex or novel legal problems still often require human legal expertise. Finally, questions of liability and regulatory oversight remain largely unresolved in many jurisdictions, creating uncertainty about accountability when AI-generated legal advice proves incorrect or harmful.

- **Case Study: ODR Platforms in British Columbia (Canada) – Civil Resolution Tribunal**

[17] Press Release, Fed. Trade Comm'n, FTC Finalizes Order with DoNotPay That Prohibits Deceptive 'AI Lawyer' Claims (Feb. 15, 2025), https://www.ftc.gov/news-events/news/press-releases/2025/02/ftc-finalizes-order-donotpay-prohibits-deceptive-ai-lawyer-claims-imposes-monetary-relief-requires.

[18] Hannes Westermann, Using Artificial Intelligence to Increase Access to Justice 45-46 (2023) (Ph.D. dissertation, Université de Montréal),

The Civil Resolution Tribunal (CRT) in British Columbia, Canada, represents one of the world's most advanced implementations of online dispute resolution with AI assistance. Established in 2016, the CRT was Canada's first online tribunal and among the pioneering examples of integrating online dispute resolution directly into a public justice system.[19]

The tribunal was initially created to resolve strata property (condominium) disputes but has since expanded to handle small civil claims under CAD 5,000, motor vehicle injury disputes, and certain society and cooperative association disputes.

The CRT's process incorporates AI assistance across multiple stages of dispute resolution. First, the platform offers a "Solution Explorer" that uses a guided pathway model with embedded AI to help users identify their legal issues and understand their rights. The system asks questions in plain language and adapts subsequent inquiries based on previous responses, effectively creating a personalised legal diagnosis.[20] This diagnostic phase helps parties clarify their legal position before formal proceedings begin.

Once a claim is filed, the platform facilitates a multi-stage resolution process beginning with party-to-party negotiation through a structured online interface. If direct negotiation fails, the system transitions to facilitated settlement with a human mediator. Throughout this process, AI tools assist by suggesting potential compromise positions based on outcomes in similar cases and by identifying misunderstandings or areas of potential agreement from party communications. If mediation proves unsuccessful, a tribunal member (adjudicator) makes a binding decision, with AI assistance in

[19] Shannon Salter, Online Dispute Resolution and Justice System Integration: British Columbia's Civil Resolution Tribunal, 34 WINDSOR Y.B. ACCESS JUST. 112 (2017).
[20] Shannon Salter & Darin Thompson, Public-Centred Civil Justice Redesign: A Case Study of the British Columbia Civil Resolution Tribunal, 3 MCGILL J. DISP. RESOL. 113 (2017).

researching relevant precedents and drafting standardised components of decisions.

6. Ethical, Legal And Regulatory Challenges

6.1. Algromathic Accountability And Transparency

The implementation of AI sentencing tools in the United States, particularly the COMPAS system, highlights critical issues of racial bias and procedural due process that must be addressed for responsible AI integration in judicial systems. As previously discussed, the ProPublica investigation of 2016 laid bare significant racial disparities in COMPAS assessments. Their analysis of over 7,000 defendants in Broward County, Florida, found that Black defendants were significantly more likely to be incorrectly classified as high-risk (44.9% false positive rate) compared to white defendants (23.5%).[21] Beyond racial bias concerns, COMPAS implementation has highlighted fundamental due process issues.

6.2. Data Protection And Privacy Concerns

The implementation of AI-driven court systems raises significant data protection and privacy concerns, particularly regarding the handling of personal information within judicial processes. Court records often contain sensitive personal data, including health information, financial details, family circumstances, and allegations of misconduct, that requires robust protection against unauthorised access, misuse, or exploitation.

In India, the evolving data protection framework creates both opportunities and challenges for judicial AI implementation. The *Digital Personal Data Protection Act, 2023,* establishes comprehensive requirements for personal data processing, including principles of purpose limitation, data minimisation,

[21] Timothy O'Brien, Compounding Injustice: The Cascading Effect of Algorithmic Bias in Risk Assessments, 13 GEO. J.L. & MOD. CRITICAL RACE PERSPS. 39 (2021).

and storage limitation that directly impact judicial data management practices. Several specific privacy concerns arise in AI-driven judicial contexts*:* ***Re-identification risks, Function creep, Algorithmic profiling, Third-party access, and Cross-border data flows.*** Addressing these concerns requires comprehensive governance frameworks that incorporate principles of privacy by design, data minimisation, and meaningful transparency. Privacy-enhancing technologies, such as differential privacy, federated learning, and homomorphic encryption, can enable beneficial AI applications while minimising privacy risks.

6.3. Legal Personhood And Decision-Making Authority

The increasing sophistication of AI systems in judicial contexts raises profound questions about their proper classification as advisory or determinative entities and the constitutional implications of replacing human discretion with algorithmic decision-making.

In contrast, the determinative approach would grant AI systems greater autonomy in certain decision-making contexts, potentially allowing them to resolve straightforward cases or procedural matters without direct human intervention.

Proponents argue that for routine, non-complex matters, fully automated decisions might enhance efficiency and consistency. Estonia's exploration of AI for small claims adjudication (though more limited than initially reported) represents a tentative step in this direction, though with significant human oversight mechanisms (Estonia Ministry of Justice, 2019).[22]

[22] Ministry of Just. & Digital Affs., Republic of Est., Estonia Does Not Develop AI Judge (Mar. 26, 2019), https://www.justdigi.ee/en/news/estonia-does-not-develop-ai-judge

The broader question of legal personhood for AI connects to these judicial application issues while extending beyond them. Legal personhood traditionally confers both rights (such as property ownership, contractual capacity, and standing to sue) and responsibilities (such as liability for harms and compliance with legal duties). Current legal systems typically treat AI systems as property owned by humans or corporate entities rather than as legal persons in their own right.

7. The Way Forward: Policy And Regulatory Recommendations

7.1. JUDICIAL GUIDELINES ON THE USE OF AI

The establishment of national AI ethics frameworks specifically tailored to judicial applications represents an essential step toward the responsible integration of artificial intelligence in court systems. These frameworks should articulate core values and principles to guide AI development, deployment, and governance in judicial contexts, while acknowledging the unique ethical considerations that arise when algorithmic systems influence fundamental rights and legal determinations.

Several jurisdictions have begun developing such frameworks. The United States National Centre for State Courts (NCSC) published *"Artificial Intelligence: Guidance for Use of AI and Generative AI in State Courts" in 2024*, articulating principles including constitutional compliance, transparency, fairness, accuracy, human oversight, and appropriate use limitations (NCSC, 2024).[23]

Similarly, the Council of Europe's European Commission for the Efficiency of Justice (CEPEJ) adopted the *"European Ethical Charter on the Use of AI*

[23] NAT'L CTR. FOR STATE CTS., ARTIFICIAL INTELLIGENCE: GUIDANCE FOR USE OF AI AND GENERATIVE AI IN STATE COURTS (2024), https://www.ncsc.org/__data/assets/pdf_file/0014/102830/ncsc-artificial-intelligence-guidelines-for-courts.pdf

in Judicial Systems," establishing principles of respect for fundamental rights, non-discrimination, quality and security, transparency, and user control.
These national frameworks should address several key domains: ***Appropriate application boundaries, Transparency requirements, Oversight mechanisms, Training and competency standards, and Procurement guidelines.***

Judicial academies play a crucial role in developing detailed protocols for AI utilisation based on these broader frameworks. As institutions dedicated to judicial education and professional development, academies can translate high-level ethical principles into practical guidance for specific court contexts and case types. The National Judicial Academy in India, for instance, has begun incorporating AI ethics and competency modules into its training curriculum for judges, helping to prepare the judiciary for technological transition while instilling appropriate caution about AI limitations.[24]

7.2. Capacity Building And Skill Development

The successful integration of AI into judicial systems requires comprehensive capacity building and skill development for judges, advocates, and court staff. As AI applications become more prevalent in legal contexts, legal professionals must develop both technical literacy and critical judgment to effectively work alongside algorithmic systems without over-reliance or inappropriate scepticism. Training programs should address several core competencies: ***AI literacy, Critical evaluation skills, Ethical judgment, and Practical usage skills.***

Capacity building efforts should particularly prioritise addressing the "black box problem" in judicial AI, ensuring that judges and other legal

[24] Shami Sumer Singh, Use of AI-Driven Support System to Help Courts in Predicting Child Custody Outcomes in India, 18 WORLD J. ADVANCED RSCH. & REVS. 2108 (2024).

professionals understand at least conceptually how the AI systems they use reach conclusions, even if they don't comprehend every technical detail. This understanding is essential for maintaining appropriate scepticism, identifying potential errors, and fulfilling explanatory obligations to parties and reviewing courts.

7.3. Encouraging Open-Source, Transparent Ai Model

The promotion of publicly accountable AI systems for judicial applications represents an essential counterbalance to proprietary, non-transparent solutions that may compromise values of openness, scrutiny, and public control fundamental to democratic legal systems. Open-source approaches offer several distinct advantages for judicial AI development, such as Transparency and scrutiny, Collaborative improvement, and Adaptability to local contexts, reduced vendor lock-in, and knowledge diffusion.

Several promising open-source judicial AI initiatives have emerged globally. The ***"OpenNYAI"*** project in India aims to develop open-source natural language processing models specifically trained on Indian legal texts in multiple languages, addressing the challenges of legal document processing in a multilingual legal system. The project has released models for judgment summarisation, precedent identification, and issue extraction that courts can freely implement and adapt.[25]

Conclusion

The potential of responsibly integrated AI to transform justice delivery is substantial, offering meaningful advances in efficiency, accuracy, and accessibility that can address persistent challenges facing judicial systems worldwide. As our analysis has demonstrated, AI applications across the

[25] S. Rohana, Transforming Legal Practice: The Rise of AI for Efficiency and Access to Justice, 4 INT'L J. LEGAL RSCH. 87 (2024).

spectrum of judicial processes, from administrative automation and legal research to risk assessment and language translation, present opportunities to enhance justice delivery while simultaneously raising important ethical, legal, and constitutional questions.

The efficiency benefits of judicial AI are perhaps most immediately apparent. Electronic case management systems, document analysis tools, and automated drafting assistance can significantly reduce administrative burdens and processing times, potentially addressing the backlog crisis that undermines timely justice in many jurisdictions. The SUPACE system in India and similar initiatives elsewhere illustrate how strategic AI deployment can free judicial resources for tasks truly requiring human judgment, potentially transforming court operations without fundamentally altering judicial roles.

GOLD, GEMS AND ALGORITHMS : REIMAGINING INDIAN JEWELLERY TRADE WITH ARTIFICIAL INTELLIGENCE

Shagoon Varma[1]

Abstract

India has been one of the first countries to make use of jewellery. Our tryst with jewellery spams from some of the first ancient necklaces to the illustrious Koh-i-noor diamond. A tradition that dates back to one of the oldest civilisation of the world, the Indus Valley Civilisation in 3300-1300 BCE, where jewellery made out of copper and gold was found. In the modern times this same interest along with mineral deposits, age old craft of jewellery making and new age technology, like Artificial Intelligence provides us with a beautiful opportunity to capitalise our craft and encourage foreign trade, by use of integrating modern and age old techniques for advancement with social accountability.

Introduction

India's legacy as a jewellery powerhouse stretches back thousands of years, where jewellery has served as both a symbol of cultural identity and a medium of economic exchange. From the intricate beadwork of the Indus Valley Civilisation to the lavish adornments of the Mughal courts, Indian jewellery reflects a blend of artistry, spirituality, and commerce. Even today, handcrafted jewellery remains an essential part of Indian traditions, while also acting as a significant economic driver in the global market.

In the more modern era, the global jewellery industry is valued at over USD 340 billion and continues to expand with growing demand in regions such as

[1] L.LM. 1st Year/ Business Group

North America, the Middle East, and Southeast Asia[1]. India plays a pivotal role, especially in the global diamond trade—from the oldest diamond refining establishment to polishing diamonds by using new age technology, however most of these diamonds are cut, polished and sent back to international markets for sale, due to which India ranks among the top exporters of gold jewellery, with key export destinations including the UAE, the USA, and Hong Kong.[2]

Jewellery exports contribute significantly to India's economy, accounting for about 7% of the national GDP and employing over 4.5 million people, particularly skilled artisans and women in semi-urban and rural areas.[3] The sector also plays a crucial role in strengthening India's foreign exchange reserves, given its position as one of the top foreign exchange earners.

This chapter aims to analyse the potential implementation of artificial intelligence for process like quality checks while incorporating the analytical skills of Ai with the skills of seasoned artisans to ensure greater optimisation and efficiency in manufacturing units for jewellery in India.

Overview of India's Jewellery Sector

India's jewellery sector is one of the most vibrant and globally integrated segments of its economy, deeply rooted in tradition while being significantly export-oriented. It encompasses a diverse range of products, each catering to different global markets.

Types of Jewellery Products Exported

The exports include handcrafted gold jewellery, diamond-studded pieces, rough and polished diamonds, coloured gemstones like emeralds, sapphires, and rubies, as well as silver ornaments and fashion or costume jewellery. India is especially dominant in the diamond segment, processing nearly 90%

[2] Statista, 2024

[3] Ministry of Commerce & Industry, 2023.

of the world's rough diamonds by volume. The country also holds a strong presence in gold and silver jewellery exports, with growing interest in affordable, semi- precious and fashion jewellery in Western and Southeast Asian markets.

Export Data & Global Share

Surveys shows that the gems and jewellery sector contributed approximately 9.2% to India's total merchandise exports in FY 2023–24, making it one of the top export categories. Major destinations for these exports include the United Arab Emirates, the United States, Hong Kong, Singapore, and several European nations. In 2023, total exports from the jewellery sector stood at USD 37.48 billion, despite a slight contraction due to global economic volatility.

Major Manufacturing and Export Hubs

These Hubs are concentrated in cities known for their specialised skills. Surat leads globally in diamond cutting and polishing, handling over 80% of the world's cut diamonds. Mumbai serves as the financial and export processing hub with almost 90% of diamonds passing it at one time or another, while Jaipur is renowned for coloured gemstone craftsmanship. Other significant hubs include Delhi, Hyderabad, and Kolkata, each with unique contributions. The sector is driven heavily by MSMEs and family-run enterprises, which dominate over 70% of the jewellery manufacturing units in the country.

Economic Significance

The jewellery industry extends beyond exports. It directly employs nearly 5 million people, particularly skilled artisans, women, and workers in informal setups. Moreover, India's heritage designs and fine craftsmanship create significant value addition, making Indian jewellery sought after globally for its aesthetic and cultural appeal.

Integration of Artificial Intelligence to Jewellery Manufacturing

Integrating AI into jewellery manufacturing is revolutionising the industry in multiple sectors. There are potentials to monetise and increase efficiency by customisation, efficiency and saving costs by implementing AI into the industry

Blockchain for Ethical Sourcing

While not a direct result, Artificial Intelligence can supplement the process of making blockchains to create provenance tracking. A blockchain is a digital decentralised system where the records with respect to manufacturing processes and life cycle of any product.

With blood diamonds and other malicious practices being a concern for illegal activities, utilisation of artificial intelligence will secure that the procurement of gemstones, precious metal and any other materials needed will remain ethical and legal. This will in turn help improve social accountability and reduce illegal trafficking.

Fraud Detection

Artificial intelligence can help us understand and analyse the activities and patterns that makes up the transactional history of a business. Based on previous experiences it is made possible to understand irregularities which in turn can be used to identify frauds.

For example, it can use a particular gemstone and find out references as to when and how many times, the particular gemstone has appeared in various entries. Activities like this can help identify and curb counterfeits and suspicious activity in real time.

Faster Quality Checks

Quality checks are strictly adhered to in manufacturing processes, this cab take hours per piece of jewellery and even then the margin for human error very much exists. Artificial intelligence offers high quality scans which can utilised to determine scratches, dents, deformities, improper finish in addition to other structural analysis. When integrated with the technology of X-Ray and Florence analysis, AI can also be made able to analyse gold purity and composition with microgram precision.

Precision Enhancement

Artificial Intelligence can make algorithms and work closely with Authoritised tools for high precision cutting, engraving and polishing. This will ensure meeting the standards of manufacturing and most efficient use of precious metals and resources utilised in the process of manufacturing.

Customisation for Consumers

Customisation and manufacturing jewellery on order only can help reduce costs for the manufacturer. As the orders can be conceptualised and customised on the wants of the customers, they need not be produced and manufactured in advance, saving cost and other resources. In addition to that artificial intelligence can be used to utilise the best cut, colour, shape and other characteristics as required, with references to current trends. This is to make a truly unique product for consumer, and with capability to render the 3D model of the same, for customers satisfaction.

Legal and Policy Framework Governing Foreign Trade in Jewellery

India’s foreign trade in jewellery operates within a multi-layered legal and policy framework that aims to promote exports while maintaining regulatory oversight. These frameworks are essential for ensuring the integrity, competitiveness, and global trust in India’s jewellery exports, especially in high-value segments like gold and diamonds.

Foreign Trade (Development and Regulation) Act, 1992

This Act serves as the cornerstone of India's trade governance system. It empowers the Central Government to develop and regulate import and export activities. A key implementing authority under this Act is the Director General of Foreign Trade (DGFT), responsible for formulating the Foreign Trade Policy, granting export licenses, and ensuring adherence to trade regulations. The Act facilitates control over sensitive sectors like gem and jewellery to prevent misuse and promote legitimate trade.

SEZs (Special Economic Zones) and Export Oriented Units (EOUs)

The SEZ and EOU frameworks offer major benefits for jewellery exporters. Units located in SEEPZ-SEZ, Mumbai—one of the oldest and largest jewellery export hubs—enjoy duty-free imports of raw materials, tax exemptions, and single-window clearances[11]. These schemes significantly reduce production costs and improve global competitiveness. EOUs are permitted to export 100% of their output, further boosting export-driven manufacturing, especially in gold and diamond jewellery.

Bureau of Indian Standards (BIS) and Hallmarking

Quality assurance is essential in the global jewellery trade. The BIS Act mandates hallmarking of gold jewellery in India, ensuring purity and transparency in both domestic and export markets. Mandatory hallmarking, aligned with international ISO standards, enhances global consumer trust and reduces the risk of counterfeit or impure products entering global supply chains[12]. Exporters must comply with BIS standards to maintain credibility in high-value international markets.

Institutional Support and Promotion

India's jewellery export ecosystem is strengthened by a network of institutions that provide financial, strategic, promotional, and risk-mitigation support to exporters. These bodies help navigate global markets and ensure the sector remains competitive on a global scale.

Gem & Jewellery Export Promotion Council (GJEPC)

Established in 1966 by the Ministry of Commerce and Industry, the GJEPC plays a crucial role in boosting India's jewellery exports by offering policy guidance, trade facilitation, and promotional activities. It organises flagship events like the India International Jewellery Show (IIJS), and facilitates buyer-seller meets across the UAE, the US, Europe, and East Asia. In 2023 alone, GJEPC hosted over 50 international B2B events, helping Indian exporters connect with over 1,500 global buyers[13]. It also runs training institutes to upskill artisans and publishes regular trade statistics and market intelligence reports that guide export strategies.

India Brand Equity Foundation (IBEF)

The IBEF, under the Ministry of Commerce, is responsible for building global awareness of Indian industries, including gems and jewellery. Through strategic marketing campaigns and participation in global trade fairs, IBEF showcases India's traditional craftsmanship and design innovation. It also leverages digital platforms and e-catalogues to brand India as a source of premium jewellery products.

EXIM Bank and ECGC (Export Credit Guarantee Corporation)

The Export-Import Bank of India (EXIM Bank) provides financial assistance and working capital loans to jewellery exporters. Meanwhile, the ECGC mitigates commercial and political risks by offering credit insurance and guarantees, ensuring that exporters can conduct business with greater confidence. In FY 2022–23, ECGC supported over 11,000 exporters, including those in the jewellery sector.

Ministry of Commerce and Industry

The Ministry of Commerce is the chief architect of India's Foreign Trade Policy (FTP) and oversees multilateral and bilateral trade agreements, such as the India-UAE Comprehensive Economic Partnership Agreement (CEPA), which has significantly boosted jewellery exports to the UAE since 2022[16]. It works in tandem with other institutions to ensure that India's export policies are aligned with global trends and domestic priorities.

5.5 Private Sector and Public Private Partnerships

Tanishq, Kalyan Jewellers and CaratLane are some of the platforms which have started using AI powered platforms for their manufacturing processes, in collaboration with companies like WIPRO and TCS. This has helped us integrate the most advanced technologies to joint research and developmental teams to provide the most advanced capabilities without having any hindrances.

Case Study 1: Natural Resources and Employment Generation

India's jewellery sector is intricately tied to its rich natural resources, particularly in regions with reserves of diamonds, gold, and semi-precious stones. These regions not only supply raw materials but also drive significant employment across the value chain.

Natural Resource-Based Jewellery Sector

India is home to several resource-rich mining regions that supply raw materials for the jewellery industry. Panna in Madhya Pradesh is the only active diamond mining site in India, operated by the National Mineral Development Corporation (NMDC), producing about 37,000 carats annually[17]. In Karnataka, the Hutti Gold Mines, one of the oldest in the world, yields approximately 1.6 tonnes of gold per year, while the defunct

Kolar Gold Fields remain historically significant[18]. Rajasthan contributes to the gemstone trade with deposits of emeralds, garnets, and agates, which support the state's gem-cutting industry. These resources fuel India's gemstone and jewellery exports while sustaining local economies.

Employment Creation through Jewellery Value Chain

The jewellery sector employs over 5 million people directly, with many more in ancillary roles[19]. Employment is concentrated in artisanal hubs such as Surat (world's largest diamond cutting and polishing center), Jaipur (known for coloured stone cutting), and West Bengal, where delicate filigree work is practiced. Notably, a large number of women artisans in rural India are involved in traditional jewellery making, beadwork, and silver craftsmanship, providing livelihood in areas with limited formal employment opportunities. Government initiatives like the Skill India Mission and training programs by GJEPC's Indian Institute of Gems & Jewellery (IIGJ) have helped upskill over 100,000 artisans in the past five years.

Government Support for Local Mining and Processing

To promote inclusive growth, the government uses funds from the District Mineral Foundation (DMF) to develop mining-affected regions by investing in education, healthcare, and artisan support[21]. Additionally, the MSME Cluster Development Programme provides infrastructure, credit, and marketing assistance to small jewellery units, especially those engaged in export activities, thus creating a sustainable economic ecosystem around resource-based jewellery production.

Recent Developments and Strategic Initiatives

India's jewellery industry has seen transformative developments in recent years, driven by strategic government initiatives, international trade agreements, and technological adoption. These changes have been instrumental in enhancing global competitiveness, encouraging self-reliance, and improving transparency in the value chain.

Make in India & Atmanirbhar Bharat

Under the umbrella of Make in India and Atmanirbhar Bharat, the government has focused on bolstering domestic manufacturing of jewellery and reducing dependence on imported raw materials. A major thrust has been given to the lab-grown diamond segment, where India has emerged as a global leader. As of 2023, India contributes nearly 15% of global lab-grown diamond production, with Surat acting as the primary hub. The government has also removed customs duty on seeds used in lab-grown diamond manufacturing in the Union Budget 2023–24 to promote domestic production.

India-UAE CEPA (2022)

The signing of the Comprehensive Economic Partnership Agreement (CEPA) between India and the UAE in 2022 marked a significant milestone. The agreement offers zero-duty access to Indian jewellery exports in the UAE, India's second-largest export destination for gems and jewellery. Post-CEPA, India's gem and jewellery exports to the UAE surged by over 20% in FY 2022–23, contributing to a total trade value of USD 5.77 billion in this sector alone[25]. Additionally, CEPA facilitates faster customs clearances and mutual recognition of hallmarking standards, reducing trade friction.

FTAs and Bilateral Agreements

India is actively negotiating Free Trade Agreements (FTAs) with key regions such as the UK and the European Union, while benefiting from existing pacts with ASEAN countries. These agreements aim to ease tariffs, provide preferential access, and open new markets for Indian jewellery products. The proposed India-UK FTA is expected to reduce import duties on key jewellery segments and enhance exports to Europe's high-demand markets[26].

Technology Integration

Technological innovation is redefining jewellery production and sales. Blockchain is increasingly used for gemstone traceability, ensuring authenticity and ethical sourcing. Moreover, AI and machine learning are being adopted for automated jewellery design and personalised virtual showroom experiences. The launch of BIS's digital hallmarking system has further streamlined certification processes, making them more transparent and efficient.

Upcoming Startups

Stylumia and Dvij AI are some of the prime examples of startups that have been introduced recently to become tools for trend analysis and development in the jewellery manufacturing sector. These startups have helped in gem authentication and presented a safe and effective mechanism to enhance productivity. These startups have worked to make tech-enabled jewellery manufacturing and exports, giving a boost to the private sector of manufacturing in India.

Comparative Global Perspective

India holds a dominant position in the global jewellery trade, particularly in the cutting and polishing of diamonds, where it accounts for over 90% of the world's supply by volume[29]. However, when compared globally, India still faces critical gaps in branding, innovation, and regulatory efficiency. The integration of Artificial Intelligence has already begun in many countries in various departments of jewellery manufacturing.

India vs China:

India leads in labour-intensive, volume-driven production, while China excels in automation, advanced technology, and design innovation. China's jewellery sector is heavily integrated with e-commerce and AI-based design tools, making it more responsive to consumer trends. China has also used AI technology for mass customisation of jewellery design, especially in the jewellery manufacturing hub situated in Shenzhen. They have mixed robotics with the same to have robotic assembly line for efficient manufacturing.

Thailand:

Recognised for its government-backed gem and jewellery institutes, Thailand has invested significantly in international-grade gem-testing laboratories and design education, enhancing product quality and global consumer trust.

Italy:

Italy focuses on heritage-based luxury branding and craftsmanship. Italian jewellery houses benefit from strong EU market access, robust artisan training ecosystems, and high-value exports in premium gold and designer jewellery. They have also integrated Artificial Intelligence into high end brands like Bulgari.

UAE:

Dubai, in particular, has emerged as a global gold trade hub, offering tax-free zones, streamlined customs processes, and advanced logistics for jewellery re-export. The implementation of the Dubai Gold and Commodities Exchange (DGCX) has added further trust and efficiency to their trade model. They have started using AI with consumer analytics and immersive retail.

Salient Observations

To remain competitive and expand its global footprint, India must focus on streamlining export compliance through efficient single-window clearance systems that reduce bureaucratic delays. Additionally, shifting the focus from generic production to brand-led jewellery exports can help create distinct identity and value in international markets. Embracing digital tools such as AI for design, blockchain for traceability, and virtual marketplaces can significantly enhance transparency, innovation, and consumer trust. The integration of Artificial Intelligence has helped open an altogether new dimension to the same and enhance the scope of jewellery manufacturing sector.

Key Challenges and Recommendations

India's jewellery export industry, contributing around 9–11% of India's total merchandise exports in FY 2023–24, plays a pivotal role in foreign trade. However, several obstacles continue to impede its growth trajectory.

Challenges

Increased Investments:

The tools needed pro incorporate Artificial Intelligence into the manufacturing industry are quite high, especially as initial investment and the costs for subsequent maintenance. This makes the technology inaccessible to small scale manufacturers.

Lack of Technical Skills and Awareness:

The widespread lack of technical knowledge and ability extends to jewellery manufacturing. The local artisans have also resisted the idea of same, for the fear of being replaced by technology.

High Import Duties on Gold:

Gold imports attract a combined duty of approximately 15% (including Basic Customs Duty and other cesses), making Indian jewellery less price-competitive globally[35].

Skill Gaps in High-End and Contemporary Jewellery Design:

Despite artisanal excellence, there is a lack of trained professionals adept at modern design trends, digital tools, and international fashion aesthetics.

Policy Suggestions

Teaching Artificial Intelligence to artists:

Incentivised jewellery companies can train youth in AI powered manufacturing design. This can be further developed by providing loans and subsidies to the MSME jewellery manufacturing industry.

Encourage Local Sourcing of Precious Materials:

Expand domestic exploration and processing of gold, diamonds, and coloured stones through public-private partnerships, reducing reliance on imports.

Strengthen Anti-Fraud Mechanisms:

Tighten regulations for Letters of Undertaking (LoUs), bank guarantees, and audit trails to prevent large-scale frauds like the Nirav Modi–PNB scam ($2 billion fraud)[37]. Integrate blockchain technology for better sourcing of raw materials.

Expand Export Incentives and Awareness for MSMEs:

Provide targeted training, access to RoDTEP benefits, and easier SEZ/EOU registration to help smaller players enter the export ecosystem.

Promote Green Mining and Sustainable Manufacturing:

Align with global ESG norms by incentivising clean technology adoption, responsible sourcing, and blockchain-enabled supply chain traceability[38].

Conclusion

The integration of Artificial Intelligence in India's jewellery industry marks a transformative shift, blending centuries-old craftsmanship with cutting-edge technology. From automated design and precision manufacturing to enhanced quality control and personalised retail experiences, AI is streamlining operations while expanding global market access. Strategic initiatives by industry leaders, startups, and government bodies reflect a growing commitment to innovation and digitalisation. However, challenges like high costs, skill gaps, and resistance to change must be addressed through inclusive policies, training, and infrastructure support. As India aspires to become a global leader in smart jewelry manufacturing and exports, a balanced approach that preserves artisanal heritage while embracing AI-driven growth will be key to sustained success.

References / Bibliography

1. Gems and Jewellery Export Promotion Council (GJEPC). (2024). Annual Export Performance Report 2023–24. (Retrieved from https://www.gjepc.org)
2. GJEPC. (2023). Gem and Jewellery Trade Update April–September 2023. (Retrieved from https://gjepc.org/pdf/ market_reports/Gem-and-Jewellery-Trade-April-Sept-2023.pdf)
3. GJEPC. (2024). Gem & Jewellery Trade Trends Q1 2024–25. (Retrieved from https://gjepc.org/pdf/ market_reports/G%26J_report_1st_quarter-2024-24.pdf)
4. Statista. (2023). India: Gems and Jewelry Export Value 2023. (Retrieved from https://www.statista.com/ statistics/624017/export-value-of-precious-gems-and-jewelry-india/)
5. Statista. (2024). India: Gems and Jewelry Export Value by Type 2024. (Retrieved from https:// www.statista.com/statistics/652781/export-value-of-gems-and-jewelry-by-type-india/)
6. Ministry of Commerce and Industry, Government of India. (2023). Trade Statistics. (Retrieved from https:// www.commerce.gov.in/trade-statistics/)
7. Directorate General of Foreign Trade (DGFT). (2023). Foreign Trade Policy 2023. (Retrieved from https:// www.dgft.gov.in)
8. India Brand Equity Foundation (IBEF). (2024). Gems and Jewellery Industry Report. (Retrieved from https:// www.ibef.org/industry/gems-jewellery-india.aspx)
9. Reserve Bank of India (RBI). (2023). Handbook of Statistics on Indian Economy. (Retrieved from https:// www.rbi.org.in)
10. Export Credit Guarantee Corporation of India (ECGC). (2023). Annual

Report 2022–23. (Retrieved from https://www.ecgc.in)

11. EXIM Bank of India. (2023). Export Finance for Indian Jewellery Sector. (Retrieved from https:// www.eximbankindia.in)
12. Bureau of Indian Standards (BIS). (2023). Hallmarking Regulations and Compliance Framework. (Retrieved from https://www.bis.gov.in)
13. United Nations Environment Programme (UNEP). (2023). Sustainable Mining and ESG Practices in the Global Supply Chain. (Retrieved from https://www.unep.org)
14. PHD Chamber of Commerce and Industry. (2023). Policy Recommendations for Jewellery Sector Export Growth. (Retrieved from https://www.phdcci.in)
15. Economic Times. (2022). India-UAE CEPA: Boost to Gems and Jewellery Exports. (Retrieved from https:// economictimes.indiatimes.com)
16. Financial Express. (2023). India’s Lab-Grown Diamond Industry Sees 25% YoY Export Growth. (Retrieved from https://www.financialexpress.com)
17. Business Standard. (2023). Nirav Modi and the PNB Scam: Lessons in Export Due Diligence. (Retrieved from https://www.business-standard.com)
18. Ministry of Skill Development and Entrepreneurship. (2023). Skill India Annual Progress Report. (Retrieved from https://www.msde.gov.in)
19. India Jewellery Review. (2024). Global Trends and Technological Innovation in Jewellery Design. (Retrieved from https://www.indiajewelleryreview.in)
20. World Gold Council. (2024). Gold Demand Trends – Q1 2024. (Retrieved from https://www.gold.org)
21. Invest India. (2024). Jewellery Sector Snapshot and Investment Trends. (Retrieved from https:// www.investindia.gov.in)
22. Comptroller and Auditor General of India (CAG). (2023). Audit Report on Export Incentive Schemes. (Retrieved from https://cag.gov.in)
23. JewelleryNet Asia. (2023). Thailand and Italy’s Jewellery Design Ecosystems: Comparative Study. (Retrieved from https://www.jewellerynet.com)
24. UNCTAD. (2023). Trade and Development Report – Regional Focus: South

Asia. (Retrieved from https:// unctad.org)

25. Reuters. (2025). India's Polished Diamond Exports Hit Two-Decade Low, Industry Group Says. (Retrieved from https://www.reuters.com/markets/commodities/indias-polished-diamond-exports-hit-two-decade-low- industry-group-says-2025-04-14/)

26. Reuters. (2025). India Weighs Export Support Measures Amid U.S. Tariff Hike. (Retrieved from https:// www.reuters.com/world/india/india-weighs-export-support-measures-amid-us-tariff-hike-say-government- sources-2025-04-09/

Bharati Law Review

Impact Factor 1.7348
(ISSN: 2278-6996) (E-ISSN: 2457-0567)
(RNI- MAHENG/2012/43568)

website - https://bharatilawreviewnlc.com
Email - blr@bharatilawreviewnlc.com

Call For Research Papers

Theme : Recent Development in Constitutional Law in India

The editorial board of Bharati Law Review (BLR) requests eminent scholars in the field of law to contribute unsolicited manuscripts of research papers, case analysis to BLR which aims to promote research work in law and related disciplines.

BLR is a refereed journal dedicated to the latest advancement in law. The goal of this journal is to promote quality dialogue on any legal issue.

BLR focus on issues related to Constitutional Law, Administrative Law, Civil Procedure Law, Criminal Law, Criminal Procedure Law, Domestic Law, Economic Law, Environmental Law, Intellectual Property Law, Private international Law, Jurisprudence, Media Law, Arbitration and Conciliation Law, E-Commerce, Banking, Insurance, Information Technology, Cyber Security, and emerging trend in allied subjects.

The distinguished legal scholarly contribution in the form of “Justice of words” will be appreciated and acknowledged by the members of the society in the years to come, and shall make them immortal.

The last date of submission is December 2025. The authors can submit the soft copy of the manuscript in M.S. word conforming the ILI – Indian Law Institute, system of citation at the e-mail address: blr@bvpnlcpune.org.

Mode of Citation

vol. BLR p. (yr.).

RNI – MAHENG/2012/43568

Bharati Vidyapeeth
(Deemed to be University), Pune, India
New Law College
Educational Complex, Erandwane,
Paud Road, Pune - 411 038, Maharashtra, India
Tel.- (020)25444616, Fax - (020)25455854
website - https://bharatilawreviewnlc.com
Email - blr@bharatilawreviewnlc.com