

“Balancing Innovation and Ethics: Implications of the EU AI Act for India’s Digital Regulation”

Siddhant Kondarear, PhD scholar Department of Law, Shri Jagadish Prasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India.

Dr. Vinod Kumar, Department of Law, Shri Jagadish Prasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India.

Abstract:

As artificial intelligence continues to reshape industries and societal functions worldwide, policymakers face the dual imperative of promoting innovation while safeguarding fundamental rights and ethical standards. The European Union’s AI Act represents a pioneering regulatory framework that seeks to address this challenge through a risk-based approach, categorizing AI systems according to their potential harm and imposing corresponding obligations. This structured taxonomy not only aims to mitigate risks associated with high-impact AI applications but also encourages responsible development by setting clear transparency requirements, such as mandatory disclosure of AI use and explainability provisions. Moreover, the AI Act’s extraterritorial scope underscores the EU’s commitment to assert regulatory influence beyond its borders, ensuring that AI systems deployed within its market adhere to its standards regardless of origin. This approach has significant implications for global AI governance, as it pressures non-EU entities to comply with stringent ethical and safety norms if they wish to access European markets. By balancing innovation incentives with robust safeguards, the Act strives to foster trust in AI technologies while setting a precedent for international regulatory cooperation and harmonization in this rapidly evolving field. By juxtaposing the European Union’s rights-centric regulatory framework with India’s currently fragmented digital governance landscape, this study highlights critical lessons for shaping AI policy. The EU’s approach, characterized by comprehensive risk assessment and strong protections for individual rights, offers a robust model for mitigating ethical and legal challenges posed by AI technologies. However, the EU’s stringent compliance requirements, while effective in safeguarding privacy and accountability, may not be entirely transferable to India’s dynamic and rapidly growing startup ecosystem, where innovation thrives on flexibility and speed. Recognizing these contextual differences, the paper advocates for a hybrid regulatory model that balances ethical imperatives with economic pragmatism. Such a model would draw selectively on the EU’s risk mitigation strategies, adapting them to India’s unique socio-economic environment, regulatory capacity and market conditions. This tailored framework aims to foster responsible AI development by embedding ethical standards and accountability mechanisms without imposing overly burdensome constraints that could hinder innovation and global competitiveness.

Keywords: **AI Governance, EU AI Act, India Digital Regulation, Algorithmic Accountability, Responsible AI, Data Privacy, Innovation Policy.**

Introduction: -

Flynn Coleman characterizes the current era as an “intelligent machine age” and asserts that human cognitive superiority is diminishing due to the advent of self-evolving machines. He cautions that, instead of implementing measures to prioritize humanitarian ideals in technological advancements, humans are preoccupied with forecasting the consequences of the AI revolution.

The AI Act serves as a fundamental regulatory framework within a wider array of strategic initiatives aimed at fostering the development and deployment of trustworthy artificial intelligence across the European Union. It establishes clear legal requirements and obligations for AI systems, ensuring that innovations align with ethical principles, safety standards and respect for fundamental rights.¹ This regulatory approach complements other key policy efforts, such as the Artificial Intelligence Innovation Package, which focuses on accelerating

¹“Artificial Intelligence’s Fair Use Crisis”, Benjamin L. W., Sobel, 41 Colum. J.L. & Arts 45 2017, Page No. 57.

AI research, development and adoption by supporting innovation ecosystems and public-private partnerships. In addition, the AI Act works in synergy with the Coordinated Plan on AI, which outlines a collaborative roadmap for EU member states to align their national AI strategies, investments and regulatory measures. Together, these initiatives create a comprehensive policy environment that balances technological advancement with societal trust and accountability. By integrating regulatory oversight with innovation support and coordination among stakeholders, the EU aims to position itself as a global leader in trustworthy AI, promoting sustainable growth while safeguarding citizen's rights and fostering public confidence in AI technologies.²

The Artificial Intelligence Act, enacted by the European Union, represents a pioneering legislative effort to regulate AI technologies comprehensively. As the first wide-ranging legal framework specifically dedicated to AI, the Act addresses the multifaceted risks associated with AI deployment, aiming to ensure that the development and application of AI systems align with fundamental human rights and ethical principles. By establishing clear requirements and obligations for both developers and users, the Act seeks to foster transparency, accountability and safety in AI technologies, thereby promoting trust and responsible innovation within the AI ecosystem. Furthermore, the Act emphasizes a human-centric approach to AI, prioritizing the protection of individuals and society from potential harms while enabling technological progress. It introduces a risk-based classification system for AI applications, mandating stricter controls and oversight for high-risk systems that could impact health, safety or fundamental rights.³ This regulatory framework not only encourages the conscientious use of AI but also sets a precedent for global AI governance by balancing innovation with robust safeguards, reflecting the EU's commitment to ethical and sustainable technological advancement.⁴

Procedural and Substantive Safeguards:

The AI Act places a strong emphasis on transparency as a fundamental mechanism to safeguard individual rights and uphold ethical standards in the deployment of AI systems. By explicitly linking transparency to the protection of fundamental rights, the Act mandates that providers must proactively disclose any foreseeable risks that AI applications may pose not only during their intended use but also in cases of reasonably foreseeable misuse. This comprehensive approach ensures that stakeholders, including users and regulators, have access to critical information that highlights potential health, safety or rights-related hazards, thereby enabling informed decision-making and accountability. Beyond risk disclosure, the AI Act introduces additional transparency requirements aimed at enhancing the explainability and fairness of AI systems. Providers are required to supply detailed information that clarifies how AI models operate, including their performance across different demographic groups and the nature of the data used in training and validation. This focus on explainability and demographic performance helps to identify and mitigate biases, promoting equitable outcomes and fostering trust in AI technologies. Together, these provisions create a robust framework where transparency serves as both a protective and enabling factor for fundamental rights in the context of AI deployment. The AI Act's Article 14 establishes a critical framework for human oversight over AI systems, emphasizing the protection of fundamental rights as well as health and safety. This oversight mechanism ensures that AI technologies do not function independently of human control or ethical considerations, thereby preventing decisions made by AI from undermining individual

² "Liability Regimes in the Age of AI: a Use-Case Driven Analysis of the Burden of Proof", David FernandezLlorca, Vicky Charisi, , Journal of Artificial Intelligence Research Vol 76 (2023) 613-644, Page No. 30.

³ "Law in the Era of Artificial Intelligence", Liane Colonna, Stanley Greenstein, Nordic Yearbook of Law and Informatics 2020–2021, ISBN 978-91-8892-964-8, Page No. 39.

⁴ "Artificial Intelligence and Law", Malluwawadu N.G, October 2019,

<https://www.researchgate.net/publication/341520298> . 2020,

dignity or rights. By mandating that AI systems be designed to allow intervention by natural persons, the Act reinforces the principle that human agency remains central in decision-making processes, particularly where AI outputs could have significant adverse effects on people. Moreover, the oversight required under Article 14(2) extends beyond the AI system's intended use to include reasonably foreseeable misuse, highlighting a proactive approach to risk management. This means that effective human supervision must be in place not only to monitor proper functioning but also to anticipate and mitigate potential harms arising from unintended applications.⁵ The emphasis on oversight being substantive rather than merely symbolic underscores the necessity for robust, practical mechanisms that ensure accountability and safeguard fundamental rights throughout the AI lifecycle.

The AI Act's introduction of the mandatory Fundamental Rights Impact Assessment (FRIA) represents a crucial procedural safeguard aimed at ensuring that high-risk AI systems are deployed in a manner consistent with fundamental human rights. Specifically mandated under Article 27, the FRIA requires deployers operating AI in sensitive domains—such as public administration, law enforcement and essential service provision—to systematically evaluate and mitigate potential adverse impacts on individuals' rights. This includes safeguarding privacy, preventing discrimination, ensuring due process and protecting access to education and social welfare. By embedding this assessment into the deployment process, the AI Act seeks to promote transparency, accountability and ethical governance in the use of AI technologies that have the capacity to significantly affect people's lives. Moreover, the FRIA framework compels organizations to proactively identify risks related to the infringement of fundamental rights before deploying AI systems, thereby facilitating early intervention and risk management. This procedural innovation not only aligns AI deployment with established human rights standards but also fosters trust among stakeholders, including the public and regulatory bodies. It underscores the AI Act's broader objective to balance technological innovation with robust protections against potential harms, ensuring that AI's benefits do not come at the expense of individual freedoms and social equity. Through the FRIA, the Act operationalizes a rights-centered approach to AI governance, emphasizing the importance of ethical considerations in high-stakes applications.

OECD Principles:

In 2019, the OECD Principles for Artificial Intelligence were formally adopted by 46 countries, including 38 OECD member states and 8 non-member countries. These principles serve as a foundational framework aimed at guiding the responsible development, deployment and use of AI technologies across various sectors and industries globally. They emphasize key values such as transparency, accountability, fairness and the protection of human rights, ensuring that AI systems are designed and implemented in ways that promote trust and mitigate risks associated with automation and algorithmic decision-making.⁶ The OECD Principles also encourage international cooperation and the sharing of best practices to foster innovation while safeguarding societal well-being. By setting common standards, these guidelines help align governments, businesses and researchers toward ethical AI advancements, addressing challenges like bias, privacy and safety.⁷ Ultimately, the principles seek to maximize the benefits of AI technologies while minimizing potential harms, supporting sustainable economic growth and inclusive social progress worldwide.

⁵ "Multidisciplinary Perspectives on Artificial Intelligence and the Law", Henrique Sousa Antunes, Pedro Miguel Freitas, ISSN 2352-1902, Law, Governance and Technology Series, Page No. 20.

⁶ "Multidisciplinary Perspectives on Artificial Intelligence and the Law", Henrique Sousa Antunes, Pedro Miguel Freitas, Law, Governance and Technology Series, Volume 58, Springer ISSN NO. 2352-1902, Page No. 28.

⁷ "Introduction to Machine Learning", Alex Smola, S.V.N. Vishwanathan, Cambridge University Press, 2008,

ISBN 0 521 82583 0, Page No. 32.

Subsequently, within the Committee of Experts (CoE), the Committee on Artificial Intelligence has been actively engaged in addressing the challenges posed by AI technologies. By the end of 2023, it presented a comprehensive proposal for a Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law. This proposal underscores the necessity of establishing international standards to ensure that AI development and deployment respect fundamental rights and democratic principles. Given the transnational nature of AI systems, which operate beyond the jurisdictional boundaries of any single state and can impact societies worldwide, the Committee emphasizes the importance of coordinated regulatory approaches at both international and regional levels. At the European level, this perspective is reflected in the European Commission's 2021 proposal for the first legal framework specifically targeting artificial intelligence, which forms part of the broader EU Data Package. This regulatory initiative represents a significant milestone in the governance of AI, aiming to create harmonized rules that foster innovation while mitigating risks associated with AI applications.⁸ The framework seeks to ensure transparency, accountability and safety in AI systems, thereby protecting citizens and upholding the EU's core values.⁹ The regulation of AI systems within the EU is thus a critical step toward establishing a robust legal environment that can serve as a model for other regions and contribute to the global discourse on AI governance.

European Union Artificial Intelligence Act:

On 21 April 2021, the European Commission introduced a comprehensive proposal aimed at establishing a harmonized regulatory framework for artificial intelligence (AI) across the European Union, known as the AI Regulation. This initiative also included a coordinated plan to align actions between the Commission and Member States, designed to foster trust in AI technologies while encouraging their development and enhancement. The proposal reflects the EU's strategic approach to balancing innovation with ethical and safety considerations, ensuring that AI systems deployed within its jurisdiction meet high standards of transparency, accountability and human oversight. Following the initial proposal, significant legislative progress was made: in December 2022, the European Council adopted a common position, referred to as the "general approach," on the AI Act, signaling broad agreement among Member States on the regulation's core principles. Subsequently, in June 2023, the European Parliament established its negotiation stance on the draft AI Act. After extensive negotiations, the Council and Parliament reached a provisional agreement on 9 December 2023, marking a critical milestone toward finalizing the AI Regulation. This agreement sets the stage for a unified legal framework that aims to both protect fundamental rights and promote the sustainable growth of AI technologies within the EU.

In 2024, the European Union's AI Act reached a significant legislative milestone with formal approval from both the Council on 21 May 2024 and the Parliament on 13 March 2024, marking a critical step toward regulating artificial intelligence within the EU. The Act is designed to provide a comprehensive legal framework addressing the risks and opportunities associated with AI technologies, ensuring that AI systems operate safely and respect fundamental rights. Following its publication in the Official Journal of the EU, the AI Act will enter into force on the twentieth day, initiating a transition period of 24 months before it becomes fully applicable. This phased implementation approach allows stakeholders, including developers, providers and users of AI systems, adequate time to adapt to the new regulatory requirements. Certain provisions within the AI Act have staggered application dates to address varying levels of urgency and risk associated with different AI technologies. For instance, AI

⁸ "Artificial Intelligence and Legal Analytics New Tools for Law Practice in The Digital Age", Kevin D. Ashley, ISBN 978-1-107-17150-3, Page No. 77.

⁹ "Freedom of Expression in Generative AI: A Snapshot of Content Policies", Jordi Calvet-Bademunt, Jacob Mchangama, February 2024, Page No. 69.

systems deemed to pose an unacceptable level of risk are subject to a ban that will take effect six months after the Act's enforcement date, reflecting the EU's commitment to proactively mitigate high-risk AI applications. Additionally, specific rules targeting General Purpose AI (GPAI) models already available in the market include a 12-month compliance deadline, ensuring that these widely used systems conform to the regulation's standards within a reasonable timeframe. This tiered timeline balances the need for regulatory rigor with practical considerations for industry readiness and technological evolution.¹⁰

Risk based approach:

The Artificial Intelligence Act draft employs a nuanced, risk-based framework that categorizes AI applications according to their potential impact and associated risks, structured as a pyramid of criticality. This approach ensures that regulatory measures are proportionate to the severity of the risk posed by each AI system. At the base of the pyramid, AI applications deemed to have negligible risk are subject to minimal legal oversight, allowing innovation and development with limited bureaucratic constraints. Conversely, AI systems identified as presenting unacceptable risks are outright prohibited to safeguard public safety, fundamental rights and ethical standards. Between these two extremes, the Act introduces a graduated set of regulatory requirements that intensify in rigor as the risk level increases. For moderate-risk applications, the framework encourages voluntary compliance through non-binding soft law instruments such as self-assessments and adherence to codes of conduct, promoting responsible AI development without heavy-handed enforcement. For high-risk AI systems, the draft mandates stringent, externally audited compliance obligations throughout the entire lifecycle of the application, ensuring continuous monitoring, transparency and accountability. This layered enforcement mechanism balances flexibility and control, fostering trust and safety while supporting technological progress.¹¹

Unacceptable Risk AI systems are classified into four main categories, two of which involve the cognitive and behavioral manipulation of individuals or specific vulnerable groups. These systems pose significant ethical and safety concerns because they can influence decision-making, emotions or actions in ways that undermine autonomy or exploit susceptibility. The remaining two prohibited categories include social scoring systems, which assess and rank individuals based on behavior or characteristics and real-time or remote biometric identification systems, which enable continuous or distant monitoring of individuals without their explicit consent. Despite these broad prohibitions, exceptions exist for each category, allowing certain uses under strict conditions. The key determinant for labeling an AI system as an Unacceptable Risk is the presence of harm—meaning the system's deployment must be likely to cause significant physical, psychological, social or economic damage. This harm criterion ensures that regulatory frameworks focus on preventing AI applications that could cause serious adverse impacts, while permitting beneficial uses that do not meet this threshold.

High Risk:

A High-Risk AI system must be developed through a rigorous, structured process that includes internal ex ante AI Impact Assessments and adherence to Codes of Conduct, which are overseen by inclusive, multidisciplinary teams to ensure comprehensive risk evaluation and ethical compliance. Once developed, the system is subject to an approved conformity assessment that verifies its compliance with the EU AI Act requirements throughout its entire lifecycle. This assessment may involve an external notified body, particularly for certain

¹⁰ "Codifying The Human Right To Science", William A. Schabas, The International Journal of Human Rights, , 2023, ISSN: 1364-2987, <https://doi.org/10.1080/13642987.2023.2269091>

¹¹ "Consumer Law and Artificial Intelligence, Challenges to the EU Consumer Law and Policy Stemming from the Business's Use of Artificial Intelligence", Final report of the ARSTY project, 28-29 June 2018, European University Institute, Florence, Page No. 60.

categories of high-risk systems, ensuring an additional layer of independent auditing. The process is dynamic, requiring repeated conformity assessments whenever significant changes are made to the system, thereby maintaining continuous benchmarking, monitoring and validation of its safety and reliability. Following successful conformity assessment, the High-Risk AI system must be registered in a dedicated EU database, providing transparency and traceability within the European market. A formal declaration of conformity must be signed and the system is required to bear the CE marking, signifying its compliance with EU regulations and readiness for market entry. This comprehensive regulatory framework ensures that High-Risk AI systems meet stringent standards of safety, accountability and performance before being deployed, thereby fostering trust and protecting users within the European Union.¹²

Limited Risk:

Limited-risk AI systems under the EU Artificial Intelligence Act are characterized by their interaction with humans or their ability to generate synthetic content, such as text, images or audio. These systems are deemed to carry a lower potential for harm compared to high-risk AI applications, which may impact fundamental rights or safety more severely. Consequently, limited-risk AI is not subject to the rigorous compliance requirements that govern high-risk systems, such as detailed risk management, conformity assessments or strict oversight mechanisms. Instead, limited-risk AI systems are primarily subject to transparency obligations as outlined in Article 50 of the Act. This means that users must be clearly informed when they are interacting with an AI system rather than a human, ensuring informed consent and awareness. This transparency requirement aims to safeguard user autonomy and trust, preventing deception or confusion in human-machine interactions.¹³ By mandating disclosure without imposing onerous compliance burdens, the regulation strikes a balance between encouraging innovation in AI technologies and protecting users' rights and safety in everyday applications.

Minimal Risk:

Under the EU Artificial Intelligence Act, AI systems classified as minimal risk are those that pose negligible or no danger to individuals' safety, livelihoods or fundamental rights. Because these systems have little potential for harm, the act imposes almost no regulatory requirements on their development or deployment. This means that developers of minimal risk AI are not subject to mandatory compliance processes, such as conformity assessments or transparency obligations, which significantly reduces the administrative and operational burden on such technologies. By exempting minimal risk AI systems from stringent oversight, the regulation aims to foster innovation and ease market entry for AI applications that are inherently safe and non-intrusive. However, this also places the responsibility on developers and users to ensure that these systems continue to operate within safe boundaries, as the regulatory framework primarily focuses on higher-risk AI that could impact fundamental rights or cause significant harm.¹⁴ This tiered approach balances the need for protection with the promotion of technological advancement.

Other Regulations which Relate AI in European Union:

The European Union has implemented additional regulations and initiatives concerning artificial intelligence (AI) that illustrate its strategy for overseeing AI development and usage.

¹² "Preserving Consumer Autonomy Through European Union Regulation of Artificial Intelligence: A Long-Term Approach", Sebastien Fassiaux, European Journal of Risk Regulation (2023), 14, Page No. 710-730.

¹³ "Preserving The Rule of Law in the Era of Artificial Intelligence", Stanley Greenstein, Artificial Intelligence and Law (2022) 30:291-323, <https://doi.org/10.1007/s10506-021-09294-4>

¹⁴ "Law and the Political Economy of AI Production", Petros Terzis, International Journal of Law and Information Technology, 2024, XX, Oxford University Press, 1-29, <https://doi.org/10.1093/ijlit/eaee001>

These measures aim to balance technological progress with safeguards for individuals and society.

The General Data Protection Regulation (GDPR):

It enforced since May 2018, represents a comprehensive legal framework established by the European Union to protect individuals' personal data. Its provisions require organizations, including those deploying AI systems, to ensure transparency in data processing activities, obtain explicit consent from data subjects and uphold individuals' rights to access, rectify or erase their personal data. This regulatory environment significantly influences how AI systems are designed and operated, particularly when they process or analyze personal information, by imposing strict obligations on data controllers and processors to safeguard privacy and data security. Moreover, GDPR's emphasis on accountability and data minimization compels AI developers to implement privacy-by-design principles and conduct data protection impact assessments when handling personal data. These requirements affect not only the technical architecture of AI systems but also their operational policies, ensuring that personal data is processed lawfully, fairly and transparently. Consequently, compliance with GDPR fosters trust among users and stakeholders by guaranteeing that AI-driven data processing respects fundamental privacy rights and mitigates risks associated with unauthorized data use or breaches.

The Digital Services Act (DSA):

It was adopted in 2022, establishes a comprehensive regulatory framework for digital platforms, particularly those leveraging artificial intelligence technologies. It aims to enhance content moderation practices by imposing stricter obligations on platforms to identify, remove or mitigate harmful content and misinformation swiftly and transparently. The DSA mandates platforms to provide clear information on their content moderation policies and algorithms, ensuring accountability and enabling users to understand how decisions affecting their online experience are made. Furthermore, the DSA introduces enhanced transparency requirements, compelling platforms to disclose data related to content dissemination and advertising, which supports regulatory oversight and empowers users. By addressing the challenges posed by the rapid spread of harmful or false information online, the DSA seeks to create a safer digital environment while balancing the protection of fundamental rights such as freedom of expression. This legislation represents a significant step toward harmonizing digital market regulations across the European Union, fostering trust and responsibility among digital service providers.¹⁵

The Digital Markets Act (DMA):

It get enacted in 2022, is a regulatory framework aimed at fostering fair competition within digital markets by specifically targeting large online platforms designated as “gatekeepers.” These gatekeepers are typically dominant companies that control access to essential digital services and infrastructure, often leveraging artificial intelligence (AI) to manage and influence user interactions and market dynamics. The DMA imposes a set of obligations and prohibitions on these platforms to prevent anti-competitive behaviors, such as self-preferencing, unfair data practices and limiting interoperability, thereby promoting a more open and contestable digital ecosystem. By addressing the unique challenges posed by gatekeeper platforms, the DMA seeks to curb their ability to exploit AI-driven control mechanisms that can stifle innovation and restrict market entry for smaller competitors. This legislation not only enhances transparency and accountability in the use of AI within these platforms but also empowers consumers and businesses by ensuring equal access and fair conditions. Ultimately, the DMA

¹⁵ “Position: Technical Research and Talent is Needed for Effective AI Governance”, AnkaReuel, Lisa Soder, Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024, Page

represents a significant step toward regulating the digital economy, balancing technological advancement with competitive fairness and consumer protection.

The EU Cyber security Strategy:

It is a comprehensive framework designed to strengthen the security and resilience of the European Union's digital infrastructure, with a particular focus on emerging technologies such as artificial intelligence (AI) systems. This strategy outlines a series of measures aimed at protecting critical digital assets from increasing cyber threats, including cyber attacks, data breaches and other malicious activities that could undermine trust and stability in digital environments. By fostering collaboration among member states, private sector stakeholders and international partners, the strategy seeks to create a unified and robust defense posture that addresses both current and future Cyber security challenges. Central to the strategy is the emphasis on ensuring the resilience of AI technologies, recognizing their growing role across multiple sectors and the unique vulnerabilities¹⁶ they may present. The EU aims to implement security standards and regulatory frameworks that promote secure AI development and deployment, mitigating risks related to data integrity, algorithmic transparency and adversarial attacks. Additionally, the strategy supports investment in research and innovation to advance Cyber security capabilities, enhance threat detection and improve incident response mechanisms. Together, these efforts contribute to building a safer digital ecosystem that supports innovation while protecting citizens, businesses and public institutions from evolving cyber risks.

India AI Governance Guidelines:

India does not currently have a single, standalone AI law. Instead, the government relies on a principle-based, "light-touch" framework anchored by the India AI Governance Guidelines and an evolving blend of existing technology, data protection and criminal laws.

India's adoption of a principle-based AI governance framework, structured around seven Sutras, represents a strategic approach to fostering innovation while ensuring safety, trust and inclusivity in artificial intelligence across various sectors. This framework serves as a foundational guideline that balances technological advancement with ethical considerations, promoting responsible AI development that aligns with societal values and legal standards. By emphasizing principles rather than prescriptive rules, the framework allows flexibility and adaptability to evolving AI technologies and diverse application contexts, thereby supporting innovation without compromising on accountability and transparency.¹⁷

The guidelines recommend the creation of several new national institutions to strengthen the governance, oversight and safety of artificial intelligence technologies. The AI Governance Group is envisioned as a central body responsible for coordinating policy development, regulatory frameworks and ethical standards related to AI deployment. This group would ensure that AI systems align with national interests, legal requirements and societal values, fostering trust and accountability in AI applications across sectors. Complementing this, the Technology & Policy Expert Committee would serve as an advisory panel of specialists who provide technical insights and policy recommendations. This committee would bridge the gap between rapidly evolving AI technologies and the regulatory environment, enabling informed decision-making and adaptive governance. Additionally, the establishment of an AI Safety

¹⁶ "Obligation for AI Systems in Healthcare: Prepare for Trouble and Make it Double?", Marinella Quaranta · Ilaria Angela Amantea, *The Review of Socionetwork Strategies* (2023) 17:275–295, <https://doi.org/10.1007/s12626-023-00145-z>

¹⁷ "Humanity's New Frontier": Human Rights Implications of Artificial Intelligence And New Technologies",

Noemi NAGY, *Hungarian Journal of Legal Studies*, Page No. 236-267, November 28, 2023.

Institute would focus on research, development and promotion of best practices to mitigate risks associated with AI, such as bias, security vulnerabilities and unintended consequences. Together, these institutions would create a robust infrastructure for responsible AI innovation and deployment at the national level.

AI governance guidelines emphasize fostering innovation rather than imposing restrictive controls, positioning artificial intelligence as a dynamic driver for inclusive economic growth and enhanced global competitiveness. By prioritizing innovation, these guidelines seek to create an enabling environment where AI technologies can flourish, unlocking new opportunities across sectors and empowering diverse stakeholders. This approach aligns with the broader national vision of Viksit Bharat 2047, which aims for a developed and technologically advanced India by the centenary of independence. The focus on innovation over restraint ensures that regulatory frameworks remain adaptive and supportive of emerging AI capabilities, encouraging investment, research and collaboration while addressing ethical and societal concerns. This balanced governance model promotes responsible AI deployment that not only accelerates technological progress but also fosters equitable development, thereby contributing to sustainable growth and the realization of India's long-term strategic objectives. India is proactively positioning itself to leverage artificial intelligence (AI) as a key driver of inclusive economic growth and enhanced global competitiveness. Recognizing AI's vast potential to revolutionize multiple sectors—from healthcare and agriculture to education and governance—the government aims to create an enabling environment that fosters innovation while ensuring equitable benefits across diverse population groups. At the same time, India is conscious of the ethical, social and security challenges that AI technologies may introduce, including risks related to privacy, bias and accountability. This dual focus underscores the need for a balanced governance approach that maximizes AI's positive impact while mitigating its potential harms to individuals and society at large. To this end, the Ministry of Electronics and Information Technology (MeitY) established a dedicated drafting committee in July 2025 tasked with formulating a comprehensive AI governance framework tailored to India's unique socio-economic context.¹⁸ The committee's mandate includes synthesizing insights from existing national laws, analyzing international regulatory trends and reviewing contemporary academic and policy research on AI governance. Furthermore, it actively incorporates public feedback to ensure that the resulting guidelines reflect the concerns and aspirations of all stakeholders. This inclusive, evidence-based approach aims to create a robust, adaptive policy framework that promotes responsible AI development and deployment in India, aligning technological progress with the country's broader development goals.

The Committee's AI governance framework is structured to provide a comprehensive and coherent approach to managing artificial intelligence in India. The first part establishes the foundational principles through seven sutras, which encapsulate the core philosophy guiding AI governance in the country. These sutras serve as ethical and strategic cornerstones, ensuring that AI development aligns with India's cultural values and socio-economic priorities. Following this philosophical grounding, the second part delves into the critical challenges and concerns associated with AI deployment, offering targeted recommendations to address issues such as transparency, accountability, inclusivity and privacy. Building on these insights, the third part outlines a detailed action plan designed to operationalize the recommendations effectively across various sectors. This plan prioritizes coordinated efforts among governmental bodies, industry stakeholders and civil society to foster innovation while mitigating risks. Finally, the fourth part translates the framework into practical guidelines aimed at industry actors and regulators, facilitating consistent and responsible implementation.

¹⁸ "Definitions of Intent Suitable For Algorithms", Hal Ashton, Artificial Intelligence and Law, Volume 31, pages 515–546, 2023

These guidelines provide a roadmap for compliance and best practices, ensuring that AI technologies contribute positively to India's development goals while safeguarding public interest.

European Union Artificial Intelligence Act and Its relevance in Indian Context:

The EU's Artificial Intelligence Act provides a comprehensive regulatory approach that emphasizes risk-based classifications, transparency, accountability, and human oversight in AI deployment.¹⁹ By adopting a similar framework, India can create a balanced legal environment that encourages innovation while safeguarding ethical standards and fundamental rights. This would enable Indian policymakers to address critical issues such as data privacy, algorithmic bias, and the potential misuse of AI technologies, which are essential for building public trust and ensuring responsible AI adoption across sectors. Furthermore, aligning India's AI regulations with international standards like those set by the EU can enhance cross-border collaboration and technology exchange. It positions India not only as a compliant market but also as a competitive player in the global AI ecosystem.²⁰ Robust and clear regulations will attract investment, foster research partnerships, and facilitate the integration of AI solutions in industries such as healthcare, agriculture, and manufacturing, ultimately driving sustainable economic growth and technological advancement.

Conclusion:

Artificial Intelligence (AI) represents a transformative technology that serves as a catalyst for innovation across all scientific disciplines, as well as within industrial and commercial sectors. The integration of AI with emerging technologies, such as Machine Learning and Cloud Computing, can expedite technological progress and enhance quality of life. However, the utilization of AI presents substantial risks to fundamental rights and values, necessitating regulation of AI system development, implementation and usage through direct risk assessment. The introduction of the AI Act and the establishment of the European AI Office signify initial steps in the regulation of AI development and application. Given the complexity of AI systems and the breadth of their application, more stringent certification standards for AI systems' safety and reliability are required than those currently proposed by the EU through the AI Act.

Establishing a comprehensive regulatory framework is essential for the responsible advancement and implementation of AI, ensuring the protection of individual rights and societal interests. This process may include the development of specific AI regulations, the enhancement of international cooperation, the improvement of technical capabilities within regulatory agencies, and the integration of ethical principles into the foundation of regulatory policies. Achieving an appropriate balance between regulation and innovation presents a considerable challenge. Excessive regulation may impede technological progress, whereas insufficient regulation could result in ethical issues and public harm. Consequently, it is crucial for policymakers to formulate regulations that foster innovation while upholding strong ethical standards.

¹⁹ "Artificial Intelligence In The Indian Criminal Justice System: Advancements, Challenges, And Ethical Implications", Aishwarya Sharma, Shivangi Chauhan Sharma, Journal of Lifestyle & SDG's Review, e-ISSN No. 2965-730X, <https://doi.org/10.47172/2965-730X.SDGsReview.v5.n01.pe04877>

²⁰ "Informational Privacy in the Age of Artificial Intelligence: A Critical Analysis of India's DPDP Act", 2023, Usha Tandon, Neeraj Kumar Gupta, Legal Issues in the Digital Age. 2025. Vol. 6, no. 2,

<https://creativecommons.org/licenses/by-sa/4.0/legalcode.en>