LAWFOYER INTERNATIONAL JOURNAL OF DOCTRINAL LEGAL RESEARCH
ISSN: 2583-7753

LawFoyer International Journal of Doctrinal Legal Research

[ISSN: 2583-7753]

Volume IV | Special Issue I

2026

Special Issue on the International E-Conference on

"Law in the Age of Artificial Intelligence: A Fundamental Perspective"

Organised by KES' Shri J. H. Patel College of Law, Mumbai, on 24 January 2026

Official Publication Partner: LawFoyer International Journal of Doctrinal Legal Research

DOI: https://doi.org/10.70183/lijdlr.2026.v4si1.41

KES' Shri. Jayantilal H. Patel Law College

About the College

KES' Shri. Jayantilal H. Patel Law College is committed to nurturing and empowering students through quality legal education. The college provides an inclusive and academically enriching environment that encourages students to excel in their studies, participate in co-curricular activities, engage in research, and develop a strong sense of civic responsibility.

The college is affiliated to the University of Mumbai, approved by the Bar Council of India, NAAC accredited with 'B++' grade, and ISO 9001:2015 certified. It offers programmes including B.A. LL.B., LL.B., LL.M., and specialised diploma courses.

With qualified and experienced faculty, a strong academic environment, moot court training, research activities and modern infrastructure, the institution seeks to prepare responsible, skilled and socially conscious legal professionals equipped to meet the challenges of a rapidly changing world.

The college functions under the aegis of Kandivli Education Society (KES), an institution established in 1936 with a long-standing commitment to quality, inclusivity and value-based education.

Affiliation	University of Mumbai
Recognition	Approved by the Bar Council of India
Accreditation	NAAC accredited 'B++'
Certification	ISO 9001:2015 certified

Kandivli Education Society

Established 1936

Kandivli Education Society (KES), established in 1936, has grown from a small educational initiative into a dynamic educational institution serving thousands of students.

With a legacy of quality, inclusivity and value-based learning, KES has expanded its academic presence across various disciplines. Its educational institutions are committed to providing meaningful learning opportunities and fostering academic, professional and personal development among students.

KES' Shri. Jayantilal H. Patel Law College is one of its prominent institutions dedicated to legal education. Through academic excellence, research, practical learning and co-curricular engagement, the college seeks to empower students with the knowledge, skills and values necessary to contribute meaningfully to society.

Educational Commitment

- Quality and inclusive education
- Value-based learning
- Academic and professional development
- Research and practical learning
- Holistic and future-ready education

Law in the Age of Artificial Intelligence

A Fundamental Perspective

About the International Conference

The International Conference on "Law in the Age of Artificial Intelligence: A Fundamental Perspective" provided an academic platform to examine the transformative impact of Artificial Intelligence on law, legal systems, governance and society.

The conference explored emerging legal and ethical questions relating to Artificial Intelligence, including intellectual property, authorship and ownership, liability, data privacy, cybersecurity, criminal justice, evidence law, human rights, ethics and accountability. It addressed the implications of Artificial Intelligence for legal education, commercial law, regulatory frameworks, policy development, research and scholarly communication.

The conference brought together academicians, research scholars, legal professionals and students to present research, exchange ideas and engage in interdisciplinary discussions. It encouraged meaningful academic dialogue and contributed towards the development of effective, ethical and inclusive legal responses to the rapidly evolving Artificial Intelligence landscape.

Conference	International Conference
Theme	Law in the Age of Artificial Intelligence: A Fundamental Perspective
Date	24 January 2026
Time	10:00 AM onwards
Mode	Online
Organised by	IQAC, Student Research Society and Library Committee
Publication Partner	LawFoyer International Journal of Doctrinal Legal Research [ISSN: 2583-7753]

Principal's Message

KES' Shri. Jayantilal H. Patel Law College

It gives me immense pleasure to welcome all academicians, research scholars, legal professionals and students to the International Conference on "Law in the Age of Artificial Intelligence: A Fundamental Perspective."

Artificial Intelligence is transforming the way societies function and is increasingly influencing legal systems, governance, professional practices and education. These developments bring significant opportunities as well as complex legal, ethical and social challenges.

This conference provided a valuable platform for intellectual exchange and interdisciplinary dialogue on emerging issues such as data privacy, intellectual property, liability, criminal justice, human rights, ethics and the governance of Artificial Intelligence. The scholarly contributions and discussions during the conference encouraged deeper understanding and promoted responsible approaches to the use of Artificial Intelligence.

I extend my best wishes to all the speakers, researchers, presenters, participants and members of the organising team for a meaningful and intellectually enriching conference.

Dr. Viral Dave
I/c. Principal
KES' Shri. Jayantilal H. Patel Law College

Organising Committee

International Conference on Law in the Age of Artificial Intelligence

I/c. Principal	Dr. Viral Dave
Convenor	Mr. Mahavir Mandot
Co-Convenors	Dr. Shweta Vijendra Pathak Ms. Priyanka Khakhadia Ms. Deepanshi Chandarana

Organised by IQAC, Student Research Society and Library Committee

KES' Shri. Jayantilal H. Patel Law College, Mumbai

Papers presented at the conference are published in the Special Issue of the LawFoyer International Journal of Doctrinal Legal Research [Volume IV, Special Issue I, 2026]

UNDERSTANDING THE FIXING LIABILITY FOR AUTONOMOUS ARTIFICIAL INTELLIGENCE: A LEGAL ANALYSIS OF DECISION-MAKING IN INTERACTIVE SYSTEMS

C. SibiKarthick[1]

I. ABSTRACT

In the international context, artificial intelligence (AI) has emerged as a transformative technological development; however, harms arising from autonomous AI activities now constitute a serious legal concern. Several international reports indicate that, following the COVID-19 pandemic, millions of children transitioned extensively to online education, resulting in increased dependence on mobile technology and, consequently, excessive engagement with social media, video games, and AI-based interactions. Interactive AI technologies, in particular, may contribute to self-harm risks among vulnerable users, especially children and adolescents. Recent incidents reveal AI-mediated harms, including AI chatbots allegedly encouraging suicide or assisting in drafting suicide notes, and voice assistants suggesting dangerous physical challenges to children. In legal parlance, one of the most sophisticated challenges lies in determining liability and identifying the responsible entity. Applying the maxim qui facit per alium facit per se, the question arises whether harm caused by an AI system can be attributed to the manufacturer, developer, company, or end user. In this regard, it becomes necessary to examine whether strict liability and product liability concepts are adequate to secure justice for victims. This paper analyses liability by examining foundational questions concerning the nature of AI, agency, personhood, intention, and justice. Using a comparative legal approach, with particular focus on the European Union model, the paper examines how risk-based regulation, mandatory human oversight, and clearly defined legal obligations may fill the existing accountability gap. It also proposes a shared but differentiated liability model to strengthen victim protection and corporate accountability. This

[1] 8th Semester B.A.LL.B Student at Government Law College, Madurai, affiliated with Tamil Nadu Dr. Ambedkar University,Tamilnadu (India).Email Id:csksibi17@gmail.com

study ultimately focuses on the problem of undefined AI liability and the significance of a legal framework to uphold justice for victims and ensure accountability.

II. KEYWORDS

Artificial Intelligence, Interactive AI, Liability, Children, Accountability.

III. INTRODUCTION

Artificial Intelligence (AI) has rapidly transitioned from a speculative technological concept to an integral component of contemporary social, economic, and governance systems. Also, it significantly supports the United Nations by promoting inclusivity, reducing inequalities, and advancing nearly 80% of the Sustainable Development Goals *(UN Report, 2026).* Its deployment in interactive platforms such as the Chatbots, Virtual assistants, recommendation sytems, and educational technologies has significantly reshaped human decision making, communication, and behaviour *(Sungho kim 2025).*

While AI-driven innovation has generated substantial societal benefits, its increasing autonomy and capacity for independent decision making have simultaneously exposed serious legal, ethical, and accountability challenges, particularly when such systems cause psychological, emotional, or physical harm. The traditional foundations of liability in law are premised upon human agency, intention, control, and foreseeability. Algorithms, especially those powered by machine learning, often operate autonomously and opaquely, making their internal functioning difficult to fully understand even for their developers *(Laxmi, & kumar, S 2025).*

This disruption complicates the attribution of responsibility when AI-mediated decisions result in grave harm, especially in sensitive contexts involving children and other vulnerable groups. Incidents involving emotionally manipulative AI interactions, encouragement of self-harm, or the suggestion of hazardous conduct underscore the inadequacy of existing liability doctrines to fully address AI induced harms. In this evolving landscape, a critical legal question emerges: who should bear responsibility for harm caused by autonomous AI system- the manufacturer, software developer, company

or the end user? The absence of clear accountability mechanisms risks creating a legal vaccum the undermines victim protection and public trust.

A. Research Objectives

The primary objective of this research is to examine the legal challenges involved in attributing liability for harms caused by autonomous and interactive AI systems, particularly where such harms affect children and adolescents. The study seeks to analyse whether existing liability frameworks, including tort liability, product liability, contractual liability, and corporate criminal liability, are adequate to address AI-mediated harm. It further aims to evaluate the relevance of the European Union's risk-based regulatory model and recent product liability developments in responding to accountability gaps. Finally, the paper proposes a shared but differentiated responsibility model that can support victim protection, corporate accountability, and restorative justice.

B. Research Questions

This paper addresses the following research questions:

1. How do autonomous and interactive AI systems challenge traditional principles of legal liability based on human agency, intention, control, and foreseeability?
2. Can existing tort, product liability, contractual liability, and corporate criminal liability doctrines adequately address AI-mediated harm, especially where minors and other vulnerable users are affected?
3. To what extent does the European Union's risk-based AI regulatory framework provide useful guidance for fixing accountability in cases involving AI chatbots, virtual assistants, and recommendation systems?
4. What form of liability model can best ensure victim compensation, child protection, and restorative justice in cases of harm caused by interactive AI systems?

 (ISSN: 2583-7753)

C. Research Methodology

This research adopts a doctrinal and comparative legal methodology. It analyses primary and secondary legal sources, including statutes, regulatory instruments, judicial materials, policy reports, academic literature, and documented case studies concerning AI-mediated harm. The doctrinal component examines the applicability and limitations of traditional liability principles, including criminal liability, tort liability, product liability, contractual liability, and corporate criminal liability.

The comparative component focuses on selected regulatory approaches in the European Union, the United States, and India, with particular emphasis on the EU Artificial Intelligence Act, the revised EU Product Liability Directive, and emerging child-safety measures relating to AI chatbots and companion technologies. Through this approach, the paper evaluates the accountability gap in existing legal frameworks and proposes a liability model directed toward prevention, compensation, and restorative justice.

IV. UNDERSTANDING THE INTERACTIVE AI SYSTEM AND ITS COMPLEXITY

A. Meaning of Artificial Intelligence and Its Types

Artificial Intelligence (AI) refers to a class of computational systems capable of performing tasks traditionally associated with human intelligence, such as learning, reasoning, perception, language processing, and decision-making (Harris, 2021). These systems rely on data-driven algorithms that enable them to operate with varying degrees of autonomy, adapt over time through learning mechanisms, and generate outputs without continuous human intervention. From a legal perspective, a formal definition of AI systems is recognised under the EU Artificial Intelligence Act *(EU AI Act 2024)* which conceptualises AI as software capable of influencing physical or virtual environments. This legal recognition is significant because it acknowledges that AI systems are no longer neutral tools but active systems capable of shaping human behaviour and decision-making.

AI systems may be classified in several ways *(IBM TEAM 2023)* however, for the purpose of this paper, two forms are particularly relevant: Generative AI and Interactive AI.

Generative AI *(Canada govt. Report 2023)* refers to systems capable of producing new content by modelling patterns learned from large datasets during training. It can create original outputs in multiple formats, including text, images, audio, and software code. Its primary function is content generation and it generally does not involve sustained real-time interaction with humans. Interactive AI *(Iterate Ai 2025)* by contrast, refers to systems designed for continuous, two-way engagement with users, which often creates the perception of human-like interaction and emotional responsiveness.

B. Nature and Functioning of Interactive AI Systems

Interactive AI systems are designed to engage directly with users rather than operating invisibly in the background. They respond dynamically to user inputs and continuously adapt their outputs based on user behaviour, contextual signals, and feedback loops *(Alsaigh, M. O., & Saeed, M. G. 2025)*. As a result, such systems significantly influence everyday decision-making by processing vast datasets to generate targeted recommendations, information, or advice *(De la Lastra, J. M. P. 2024)*.

Interactive AI systems broadly categorised into following types: conversational AI, which simulates human dialogue (for example, ChatGPT or Google Assistant or character ai); voice-based assistants, such as Alexa and Siri, adaptive recommendation systems, which personalise content and influence user choices on digital platforms *(Infobip 2025)* The functioning of interactive AI involves several interconnected stagestha is receiving user input, analysing that input using techniques such as natural language processing and machine learning, making decisions through rule-based or probabilistic models, and producing real-time outputs.

C. Problems Associated with Interactive AI: Impact on Children and Adolescents

Following the COVID-19 pandemic, children and adolescents have experienced a substantial increase in digital engagement through smartphones, online education platforms, social media, and video games *(Bergmann, C 2022)*. Alongside this shift, there has been a growing reliance on AI-enabled interactive systems such as chatbots, virtual assistants, and algorithmic recommendation tools.

This increased exposure has generated serious psychological and emotional risks, particularly because young users often lack awareness of how these systems function. Instead, they may interpret AI interactions as human-like relationships built on trust and emotional understanding *(Eufo Andoh, 2025).*

Several tragic incidents indicate the potential dangers of such reliance, although the legal allegations in these matters remain subject to judicial determination. In one pending wrongful-death action, the parents of 16-year-old Adam Raine allege that prolonged interactions with ChatGPT contributed to his death by suicide; the suit was filed in 2025 as *Raine v. OpenAI, Inc*[2], Case No. CGC-25-628528, before the Superior Court of California, County of San Francisco. The allegations, including claims concerning the adequacy of the chatbot's self-harm safeguards and redirection mechanisms, remain contested by OpenAI and should therefore be treated as pleadings rather than established findings. In another case, the mother of a 14-year-old user has alleged that prolonged interactions with Character.AI chatbots contributed to her son's death following emotionally dependent engagement with simulated human-like personalities.

AI-enabled virtual assistants also become danger. In 2021, a ten-year-old child was advised by Alexa to insert a coin into a live electrical socket as a challenge *(BBC News 2021)* Similarly, AI-driven recommendation *(Amplyfi Blog 2025)* systems on social media platforms are engineered to maximise engagement, often reinforcing addictive usage patterns and exposing children to harmful content. Real documented cases involving Instagram and TikTok demonstrate how algorithmic amplification has contributed to eating disorders, anxiety, and self-harm among vulnerable groups *(Lavi, M. 2024).*

In another significant case, Megan Garcia, individually and as personal representative of the estate of her son, Sewell Setzer III, filed *Garcia v. Character Technologies, Inc.*, No. 6:24-cv-01903, before the United States District Court for the Middle District of Florida in October 2024. The complaint alleged that prolonged interactions with Character.AI

[2] *Raine v. OpenAI, Inc.*, Case No. CGC-25-628528, Superior Court of California, County of San Francisco, Complaint filed Aug. 26, 2025.

chatbots contributed to the death of her 14-year-old son after he developed emotionally dependent engagement with AI-generated personalities.

In January 2026, Character Technologies, Google, and related defendants agreed to settle the case along with other chatbot-harm suits, and the court issued a settlement order on January 7, 2026. Reports indicate that the settlement included commitments to introduce new safety features for users under the age of 18. This development is directly relevant to AI liability discourse because it demonstrates that corporate-level legal proceedings can produce concrete safety reforms even without a final adjudication of liability.

These realincidents make it evident that interactive AI systems are no longer neutral technologies but a tool that capable of shaping behaviour and emotional states. The scale and severity of these harms expose fundamental shortcomings in existing legal liability frameworks, necessitating a deeper examination of accountability.

V. LEGAL ANALYSIS ON THE LEGAL ACCOUNTABILITY OF HARM IN THE AI ERA

A legal accountability *(Black's Law Dictionary, 10th ed. 2014)* denotes the obligation of an individual or entity to explain, justify, and bear legal consequences for actions or omissions that result in harm. Traditional accountability frameworks presuppose a human decisionmaker capable of intention, control, and moral judgment. But when It comes to AI systems fundamentally challenges this assumption by introducing algorithmic decision making with minimal or indirect human intervention.

A. Traditional Legal Liabilities

Tort and product liability offer comparatively more viable frameworks for addressing AI-mediated harm. Under negligence principles, liability may arise where a developer, deployer, or platform operator fails to exercise reasonable care in designing, testing, monitoring, or warning users about foreseeable risks associated with an AI system. In the context of interactive AI, negligence may be argued where companies deploy emotionally responsive chatbots or recommendation systems without adequate safeguards for minors, despite being aware that such systems may intensify dependency, self-harm

ideation, or exposure to dangerous content. The central inquiry is whether the corporate actor could reasonably foresee the risk and whether sufficient precautions, such as human oversight, crisis redirection, age-appropriate design, or content moderation, were adopted.

Product liability is also relevant because AI systems may increasingly be treated as products where they are placed in the market as software, applications, chatbots, virtual assistants, or algorithmic recommendation tools. The revised EU Product Liability Directive expressly recognises software, including AI systems, within the modern product-liability framework and permits compensation claims for harm caused by defective products without requiring proof of manufacturer fault.

This development is significant because victims of AI harm often face serious evidentiary difficulties in proving fault due to the opaque and adaptive nature of algorithmic systems. However, the application of product liability to AI remains complex because an AI system may operate as technically designed yet still generate harmful outputs due to defective training data, unsafe optimisation goals, inadequate guardrails, or emergent behaviour after deployment. Therefore, product liability in the AI context must assess not only manufacturing defects but also design defects, inadequate warnings, failure to update, and failure to prevent foreseeable misuse.

Additionally, the adaptive and self-learning nature of AI undermines strict liability, which assumes static risks identifiable at manufacture.

Contractual liability *(EU Expert group report 2019)* presents additional concerns. AI platforms frequently rely on terms of service and liability disclaimers to shift risk onto users and deny enforceable assurances of reliability *(Pratt, M. G. 2014).* While such terms are generally valid under private law, their legitimacy is weakened where AI systems significantly influence cognition, emotional well-being, or behaviour among minors and harms to them. Where AI systems exert substantial behavioural influence or amplify harmful content, liability cannot be excluded through contractual consent.

B. Legal Challenges in Assigning accountability

Several structural challenges prevent effective accountability for AI-mediated harm. First, definitional ambiguity surrounding concepts such as autonomy, intelligence, machine learning, and algorithmic decision-making may lead to inconsistent legal interpretation. Unlike conventional tools, interactive AI systems do not merely execute fixed commands; they generate outputs by processing data, user behaviour, probabilistic patterns, and contextual signals. This makes it difficult for courts to determine whether the harm resulted from a product defect, negligent design, inadequate supervision, user misuse, or autonomous system behaviour.

Secondly, AI systems lack legal personality, consciousness, and moral agency. They cannot form intention, knowledge, recklessness, or negligence in the legal sense. Therefore, responsibility cannot be directly attributed to the AI system itself. The law must instead trace responsibility to human or corporate actors who designed, trained, deployed, maintained, or commercially benefited from the system. This becomes particularly difficult where multiple actors are involved, including model developers, data providers, application developers, platform operators, advertisers, and end users.

Thirdly, AI decision-making is often affected by the "black box" problem. Where the internal logic of an AI system cannot be clearly explained, victims may struggle to prove causation, defect, or foreseeability. This problem is acute in cases involving psychological or emotional harm, where the causal connection between AI interaction and injury may be contested. The evidentiary burden on victims is therefore significantly higher than in traditional product or tort claims. For this reason, legal systems should consider presumptions of defect, disclosure obligations, and burden-shifting mechanisms where claimants establish that an AI system plausibly contributed to harm.

Further, AI decision-making is often opaque, giving rise to the "black box" problem, which restricts courts and regulators from tracing causation, assessing fault, and assigning responsibility *(Preksha jayaswal, IJLR 2025)*. The absence of judicial precedents and AI-specific statutory frameworks compounds this difficulty. Finally, the cross-border

development and deployment of AI systems generate jurisdictional conflicts due to divergent national regulations.

C. International Regulatory Gaps

Globally, regulatory approaches to fixing AI liability remain fragmented. The European Union has adopted one of the most structured regulatory models through the EU Artificial Intelligence Act, 2024, which follows a risk-based approach and imposes obligations relating to transparency, safety, governance, and human oversight. However, the EU AI Act primarily functions as a regulatory compliance framework and does not, by itself, constitute a complete liability regime for compensating victims of AI-mediated harm. Its importance lies in the fact that statutory duties relating to risk management, documentation, transparency, and oversight may assist courts and regulators in assessing whether AI providers or deployers acted responsibly.

The United States, by contrast, continues to rely largely on sector-specific laws, tort principles, consumer protection standards, product liability doctrines, and litigation-based accountability. This fragmented approach allows courts to adapt existing doctrines to new forms of AI harm, but it also creates uncertainty because the legal consequences of chatbot harms, algorithmic recommendations, and emotionally manipulative AI interactions may vary across jurisdictions and factual settings. Pending and settled chatbot-related litigation in the United States demonstrates the increasing relevance of corporate accountability, but the absence of a unified AI liability statute continues to create difficulties for victims seeking predictable remedies.

India's position requires a more nuanced assessment than merely referring to the Consumer Protection Act, 2019 and the Information Technology Act, 2000. While these statutes remain relevant, India has recently moved toward a broader governance framework for AI and data-driven technologies. The Digital Personal Data Protection Act, 2023 is directly relevant to interactive AI systems because such systems often depend on the collection, processing, profiling, and behavioural analysis of users' digital personal data. In the context of children, the Act is particularly significant because it requires verifiable parental consent before processing children's personal data and prohibits

processing that is likely to cause detrimental effects to the well-being of a child. This has clear implications for AI chatbots, recommendation systems, educational technologies, and social platforms that personalise content or generate emotionally responsive outputs for minors.

Further, the Reserve Bank of India's Committee to develop a Framework for Responsible and Ethical Enablement of Artificial Intelligence in the Financial Sector submitted its FREE-AI Report on August 13, 2025. Although sector-specific, the report is important because it recognises that AI governance must encourage innovation while simultaneously mitigating risks. Its focus on responsible deployment, ethical enablement, risk governance, accountability, and institutional oversight provides a useful model for regulated sectors where AI decisions may affect consumers, borrowers, investors, and other vulnerable users. The FREE-AI framework therefore illustrates India's emerging tendency to regulate AI through sectoral governance mechanisms rather than through a single comprehensive AI liability statute.

Most importantly, the Ministry of Electronics and Information Technology unveiled the India AI Governance Guidelines on November 5, 2025 under the IndiaAI Mission. These guidelines adopt a principle-based and non-statutory approach, with "Do No Harm" as a central guiding principle. They emphasise safe, inclusive, and responsible AI adoption, risk mitigation, innovation sandboxes, and a flexible techno-legal governance ecosystem. In this respect, India's model may be contrasted with the European Union's binding risk-based framework. While the EU model relies more heavily on enforceable regulatory classification and compliance duties, India's current approach favours soft-law guidance, sectoral regulation, institutional coordination, and adaptive governance.

Despite these developments, significant gaps remain. The DPDP Act addresses data protection but does not comprehensively resolve liability for psychological, emotional, or behavioural harms caused by AI systems. The FREE-AI Report is confined to the financial sector and does not directly govern companion chatbots or child-facing recommendation systems. The MeitY AI Governance Guidelines provide valuable principles but, being non-statutory, may not by themselves create enforceable rights to

compensation. Accordingly, India's emerging AI framework is important but incomplete. A dedicated liability model is still required to address opacity, autonomy, distributed responsibility, child protection, and cross-border AI harms. Its importance lies in identifying risk-based duties that may later inform liability analysis where a provider or deployer fails to comply with legally recognised safety obligations. In addition, the revised EU Product Liability Directive strengthens victim protection by bringing software and AI systems within the product-liability framework and by modernising rules relating to defective digital products.

In the United States, AI governance remains comparatively sector-specific and litigation-driven. Liability questions are generally addressed through existing tort, consumer protection, product liability, privacy, and platform-governance principles rather than through a single comprehensive AI liability statute. This approach allows courts to adapt existing doctrines to emerging harms, but it also creates uncertainty because standards may vary across states, sectors, and factual contexts. In cases involving AI chatbots or recommendation systems, claimants may face challenges in proving duty, breach, causation, defect, and foreseeability, particularly where defendants rely on contractual disclaimers, platform immunity arguments, or the technical complexity of algorithmic outputs.

India presently addresses AI-related risks indirectly through general legal frameworks such as the Consumer Protection Act, 2019, the Information Technology Act, 2000, intermediary liability rules, data protection principles, and constitutional commitments to life, dignity, and child protection. However, these laws were not specifically designed to address autonomous AI systems that interact emotionally with children, generate harmful advice, or amplify dangerous content. As a result, Indian law requires clearer statutory duties for AI developers and platform operators, particularly in relation to transparency, age-appropriate safeguards, risk assessment, grievance redressal, and compensation for victims. A comparative reading of the EU, US, and Indian approaches demonstrates that the central regulatory gap is not the absence of all legal remedies, but

 (ISSN: 2583-7753)

the absence of a clear, AI-specific liability model capable of addressing opacity, autonomy, distributed responsibility, and cross-border deployment.

VI. FIXING THE LIABILTY TOWARD RESTORATIVE JUSTICE:

A. AI companies should be responsible entity

Despite all complexity, liability for unlawful conduct and harm caused by AI chatbots, virtual assistants, and algorithmic recommendation systems should primarily rest with corporate entities *(MacCarthy, M. 2025)* that design, deploy, and commercially benefit from these technologies. These companies exercise the highest degree of control over system architecture, training datasets, deployment environments, and safety mechanisms. Consequently, AI companies are best positioned to foresee, prevent, and mitigate risks, especially when AI systems interact with vulnerable users such as children and adolescents. Anchoring liability at the corporate level also aligns with restorative justice objectives by ensuring victim compensation, behavioural reform, and meaningful accountability.

The January 2026 settlement in *Garcia v. Character Technologies, Inc.* illustrates this practical function of corporate liability: although the case did not culminate in a final judicial finding of liability, the proceedings reportedly resulted in safety-feature commitments for users under 18. Such developments demonstrate that corporate accountability mechanisms can operate not only as compensatory tools but also as preventive instruments capable of changing platform design and child-safety practices.

This approach finds strong justification in the doctrine of corporate criminal liability *(Wadhawan, M. 2025)*, which allows companies to be held criminally responsible for harms arising from organisational policies, structural failures, or systemic negligence. In the AI context, harm often results not from isolated technical errors but from inadequate risk assessments, insufficient human oversight, or profit driven deployment decisions that prioritise engagement over safety. When corporations deploy AI systems in environments accessible to minors without adequate safeguard, the resulting harm reflects organisational fault rather than autonomous machine behaviour. Accordingly,

corporations should not be permitted to evade responsibility by invoking algorithmic opacity or claims of technological autonomy.

Product liability framework further strengthens this accountability model. The revised Product Liability Directive 2024 *(EU Liability directive 2024)* marks a significant development by explicitly recognising AI systems including chatbots, virtual assistants, and recommendation tools as products, even when delivered through software or digital services. Crucially, the Directive removes the burden on victims to prove fault and instead imposes strict liability where an AI system cause harm. This shift is particularly important given the opaque, adaptive, and self-learning nature of AI, which often makes fault-based claims practically unworkable. By prioritising victim compensation and risk internalisation, strict liability offers an effective remedy for AI induced harm.

B. Preventive Regulatory Measures for Child Protection

However, liability mechanisms alone are insufficient, particularly where children and adolescents are concerned. Preventive regulation must therefore remain the primary objective. Mandatory child-safety protocols should be imposed on AI systems likely to be accessed by minors, including real-time detection of self-harm ideation, automatic referral to crisis-support services, transparent disclosure that users are interacting with AI, recurring break reminders, and prohibitions on sexual, exploitative, or psychologically harmful interactions. California Senate Bill 243, signed into law on October 13, 2025 and effective from January 1, 2026, illustrates that targeted safeguards for companion chatbots can be implemented through operative legislation without imposing a blanket prohibition on youth access to AI technologies.

In addition, age-appropriate design standards and proportionate age-verification mechanisms should be adopted, particularly for emotionally responsive or companion chatbots. New York has already enacted a directly relevant AI Companion Models statute under General Business Law Article 47, §§ 1700–1704, effective November 5, 2025. The statute requires operators of qualifying AI companions to implement safety protocols for suicidal ideation or self-harm and to notify users at the beginning of an interaction, and

periodically during extended use, that they are interacting with AI rather than a human being.

Alongside this operative framework, pending proposals such as New York S 5668 remain relevant because they seek to address liability for misleading, incorrect, contradictory, or harmful chatbot outputs. However, excessive age-gating may threaten freedom of expression and access to beneficial AI tools; therefore, child-protection measures should be calibrated, proportionate, and focused on risk-sensitive safeguards rather than absolute exclusion.

Ultimately, an effective liability regime must adopt a shared but differentiated responsibility model. Primary liability should rest with AI companies and platform operators, while secondary liability may attach to developers and deployers whose design choices materially contribute to harm. finally, international coordination through multilateral agreements focused on children's rights in digital environments is necessary to address jurisdictional fragmentation and ensure consistent global standards.

VII. SUGGESTIONS AND RECOMMENDATIONS

In order to address the accountability gap created by autonomous and interactive AI systems, this paper recommends the adoption of a shared but differentiated liability framework. Under this model, primary responsibility should rest with AI companies and platform operators that design, deploy, and commercially benefit from AI chatbots, virtual assistants, and recommendation systems. Secondary responsibility may attach to developers, deployers, data providers, and other intermediaries where their design choices, deployment practices, or omissions materially contribute to harm. Such a model would reflect the actual distribution of control within the AI ecosystem and would prevent corporate actors from avoiding liability by relying on algorithmic opacity or technological autonomy.

Secondly, mandatory child-safety protocols should be incorporated into AI governance frameworks, particularly for systems that are accessible to minors or designed to simulate emotional companionship. These safeguards should include real-time detection of self-

harm ideation, crisis-support redirection, clear disclosure that the user is interacting with an AI system, restrictions on manipulative engagement techniques, and prohibitions on sexual, exploitative, or psychologically harmful interactions with children. Age-appropriate design standards and proportionate parental-consent mechanisms may also be introduced where AI systems create foreseeable emotional or behavioural risks for minors.

Thirdly, product liability laws should expressly recognise AI systems, including software-based chatbots, virtual assistants, and algorithmic recommendation tools, as products capable of causing compensable harm. A strict liability approach should be preferred in cases involving defective or unsafe AI systems, particularly where victims cannot reasonably prove fault due to the opaque and adaptive nature of algorithmic decision-making. This would shift the burden away from vulnerable users and ensure that the risks of AI deployment are internalised by those best positioned to prevent harm.

Finally, international cooperation is necessary to address cross-border AI harms. Since AI systems are frequently developed, trained, deployed, and accessed across different jurisdictions, fragmented national regulation may leave victims without effective remedies. Harmonised standards on transparency, child protection, risk assessment, human oversight, and victim compensation should therefore be developed through multilateral cooperation. Such an approach would strengthen restorative justice by ensuring that victims receive meaningful remedies while also promoting responsible innovation and public trust in AI technologies.

VIII. CONCLUSION

This study underscores that AI chatbots, virtual assistants, and recommendation systems have evolved into influential socio-technical systems capable of causing serious harm, particularly to children and adolescents. The analysis demonstrates that traditional liability doctrines, which are largely based on human agency, intention, control, and foreseeability, are not fully adequate to address harms produced by autonomous and

 (ISSN: 2583-7753)

opaque AI systems. The difficulty of identifying fault, proving causation, and locating responsibility among multiple actors creates a significant accountability gap.

The paper therefore concludes that liability for AI-mediated harm must be fixed in a manner that reflects control, benefit, and risk creation. Corporate entities that design, deploy, and profit from interactive AI systems should bear primary responsibility, while other actors may incur responsibility according to their contribution to the harm. Ultimately, the legitimacy of artificial intelligence depends not merely on technological sophistication, but on the existence of clear legal responsibility when harm occurs. Responsible AI governance must ensure that innovation advances human well-being, trust, dignity, and justice rather than undermining them.

IX. REFERNCES

1. Bergmann, C., et al. (2022). Young children's screen time during the first COVID-19 lockdown in 12 countries. Scientific Reports, 12(1), Article 2015. https://pure.mpg.de/rest/items/item_3328631_2/component/file_3328632/content
2. Eufo Andoh, (2025, October 1). Technology and youth friendships: Many teens are turning to AI chatbots for friendship and emotional support. Monitor on Psychology. American Psychological Association. https://www.apa.org/monitor/2025/10/technology-youth-friendships
3. BBC News. (2021). Alexa tells 10-year-old girl to touch live plug with penny. BBC News. https://www.bbc.com/news/technology%E2%80%9159810383
4. Amplyfi. (2025). How AI personalisation affects our psychological wellbeing. Amplyfi Blog. https://amplyfi.com/blog/how-ai-personalisation-affects-our-psychological-wellbeing
5. Lavi, M. (2024). Targeting children: Liability for algorithmic recommendations. American University Law Review, 73, 1367–1465. https://aulawreview.org/wp-content/uploads/2024/09/Lavi.to_.Printer.pdf
6. Legal liability. (n.d.). In Wikipedia.2026, https://en.wikipedia.org/wiki/Legal_liability

7. Expert Group on Liability and New Technologies – New Technologies Formation. (2019). Liability for artificial intelligence and other emerging digital technologies (Expert Group report, European Union). European Parliament. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf
8. Pratt, M. G. (2014). Disclaimers of contractual liability and voluntary obligations. Osgoode Hall Law Journal, 51(3), 767–782. https://digitalcommons.osgoode.yorku.ca/cgi/viewcontent.cgi?article=2755&context=ohlj
9. Preksha jayaswal, The Black Box of AI: Who's to Blame? Indian Journal Of Legal Review (IJLR), 5 (5) OF 2025, PG. 901-905, APIS – 3920 – 0001 & ISSN - 2583-2344 https://ijlr.iledu.in/wp-content/uploads/2025/04/V5I593.pdf
10. MacCarthy, M. (2025, August 20). AI companies should be liable for the illegal conduct of AI chatbots. https://www.techpolicy.press/ai-companies-should-be-liable-for-the-illegal-conduct-of-ai-chatbots/
11. Wadhawan, M. (2025). Evolving legal liability norms: A discussion on the intersection of corporate criminal liability (CCL) and legal liability of artificial intelligence (AI). Indian Journal of Law and Legal Research, 7(3), 7873–788 https://3fdef50c-add3-4615-a675a91741bcb5c0.usrfiles.com/ugd/3fdef5_607f36ede12949f3a069b167d952f0ae.pdf
12. European Parliament and Council of the European Union, Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC (Official Journal L 2024/2853, 18 Nov. 2024), EUR-Lex, https://eur-lex.europa.eu/eli/dir/2024/2853/oj/eng