Access to Justice in Eastern Europe
ISSN 2663-0575 (Print)
ISSN 2663-0583 (Online)
http://ajee-journal.com

OPINION ARTICLE

Opinion Article

CHATGPT IN THE DOCK: REFLECTIONS ON THE FUTURE OF CRIMINAL LIABILITY

Raed S A Faqir

DOI:
https://doi.org/10.33327/AJEE-18-8.S-o000156

Date of submission: 19 Aug 2025
Date of acceptance: 30 Oct 2025
Date of Publication: 30 Dec 2025

Disclaimer:

The author declares that his opinion and views expressed in this manuscript are free of any impact of any organizations.

ABSTRACT

Background: *This study examines the legal challenges posed by generative AI. It highlights the limitations of traditional criminal liability frameworks in addressing harm caused by AI outputs. The research explores new models of liability to ensure accountability while protecting individual rights in the age of intelligent machines.*

Generative AI, exemplified by ChatGPT, has evolved from a mere computational tool into a cognitive agent capable of content creation, problem-solving, and decision-making. This evolution challenges traditional criminal law frameworks, raising complex questions about the attribution of Liability when AI-generated outputs result in harm or criminal conduct. The study explores these dilemmas, focusing on the shortcomings of conventional concepts of criminal liability and exploring the need for new legal paradigms.

Methods: *The research employs a descriptive-analytical and comparative methodology. It analyses national and international legislation, legal principles, and contemporary jurisprudence, with a focus on the European Artificial Intelligence Act (2024) as a model. The study examines AI's*

autonomous capabilities, the opacity of algorithmic decision-making, and the challenges of establishing causal links between AI actions and resulting harms. Case studies are used to explore potential liability models, including preventive liability and the concept of an "artificial actor."

Results and Conclusions: *The study finds that traditional frameworks of criminal accountability are inadequate for AI systems like ChatGPT, given their partial autonomy and algorithmic complexity. It highlights the potential for expanding liability to developers, operators, and users, and the necessity of flexible legal models that combine preventive, administrative, and criminal measures. The research underscores the importance of integrating legal innovation with technological oversight to safeguard individual rights while maintaining the deterrent and protective functions of criminal law.*

1 INTRODUCTION

Generative Artificial Intelligence, led today by models such as ChatGPT, represents a pivotal milestone in the trajectory of global technological transformation. It is no longer merely an executive tool but has become a cognitive agent actively participating in content creation, problem-solving, and decision-making. This shift is clear in the growing reliance on AI systems across multiple fields—including education, healthcare and law—posing unprecedented challenges to legal systems, particularly the criminal justice system. The system now faces uncertainty in identifying the "actor" in cases involving harm or criminalised conduct resulting from AI intervention.

From this standpoint, the present study, titled *ChatGPT in the Dock: Reflections on the Future of Criminal Liability*, aims to analyse the legal dilemmas posed by this model, one of the most prominent and controversial manifestations of generative artificial intelligence. Its primary objectives are: (1) to expose the shortcomings of traditional concepts of criminal liability in light of AI intervention; (2) to analyse the complex structure of liability resulting from the actions and outputs of systems like ChatGPT; and (3) to propose restructuring legal attribution rules in a manner that ensures the protection of individual rights and the effectiveness of criminal deterrence in the age of intelligent machines.

Upon AI autonomy, the study takes a two-pronged approach: theoretically, it examines criminal concepts such as intent, causation, and actor liability; practically, it examines comparative legal models, particularly the 2024 European AI Act, and their applicability to cases involving harmful or deceptive AI use, like ChatGPT. The study is grounded in a central hypothesis: current frameworks of criminal accountability are incapable of comprehending the outputs of partially autonomous non-human entities. This necessitates the development of more flexible legal models, such as collective or virtual liability, or even the establishment of a new concept: the "artificial actor."

The core legal questions this study addresses include: Who bears criminal liability when ChatGPT use results in harmful or criminal outputs? Should liability be limited to the end user, or should it extend to developers, operators, and designers? To what extent can ChatGPT be considered an independent actor contributing to the criminal outcome? What is the optimal legislative pathway to bridge the legal gap caused by the ambiguity of intent and discernment in AI-generated actions? These questions are not merely theoretical; they lie at the heart of the challenges facing criminal justice in the coming decades. Exploring them constitutes a legal and strategic necessity for ensuring the sustainability of the judicial system in a rapidly evolving digital environment.

This study on criminal liability for AI systems like ChatGPT addresses the shortcomings of existing legal frameworks, identifies the parties liable, and the challenges in assessing intent. It examines legitimate risks, regulatory gaps, and the difficulties of applying traditional criminal law to autonomous AI behaviour. Finally, it proposes future-oriented solutions to change liability frameworks, such as corporate responsibility and systemic accountability.

2 LITERATURE REVIEW

The scientific literature indicates that criminal accountability for artificial intelligence systems, particularly complex software such as ChatGPT, faces fundamental challenges related to the nature of algorithmic actions and the opacity of decision-making processes.[1] Many researchers point out that AI operates through intricate algorithmic networks beyond human control, making it more difficult to attribute any resulting harm to conventional ideas of fault or intent.[2] Traditional criminal liability, which is based on direct human action, is being reexamined amid the crisis of causal attribution.[3]

In the same context, recent legal studies have addressed the crisis of lack of control over the technical risks posed by AI, illustrating that traditional laws are incapable of regulating these new technological risks within conventional liability frameworks.[4] To

1 Alejo José G Sison and others, 'ChatGPT: More than a "Weapon of Mass Deception" Ethical Challenges and Responses from the Human-Centered Artificial Intelligence (HCAI) Perspective' (2024) 40(17) International Journal of Human-Computer Interaction 4853. doi:10.1080/10447318.2023.2225931.

2 PR Biju and O Gayathri, 'Algorithmic Solutions, Subjectivity and Decision Errors: A Study of AI Accountability' (2025) 27(5) Digital Policy, Regulation and Governance 523. doi:10.1108/DPRG-05-2024-0090.

3 Marcelo Ferrante, 'Causation in Criminal Responsibility' (2008) 11(3) New Criminal Law Review 470. doi:10.1525/nclr.2008.11.3.470.

4 Benjamin Cheatham, Kia Javanmardian and Hamid Samandari, 'Confronting the Risks of Artificial Intelligence' (2019) 2 McKinsey Quarterly 8 <https://www.mckinsey.com/capabilities/quantumblack/our-insights/confronting-the-risks-of-artificial-intelligence> accessed 10 August 2025.

address risks posed by software like ChatGPT, the legal landscape is moving toward preventive and abstract liability models, emphasising producers' accountability for technical safety measures rather than for direct harm.[5]

Moreover, the literature highlights legislative and practical challenges stemming from the exclusion of intelligent software from traditional product laws, particularly those that define a product solely as a physical entity, thereby obstructing effective accountability for such systems.[6] In this regard, the European Artificial Intelligence Act emerges as an advanced model imposing deterrent sanctions on violations and promoting enhanced transparency and disclosure of AI-related risks.[7] Recent studies also recommend integrating administrative oversight with criminal liability, thereby opening new horizons for managing technological risks through an integrated and evolving legal framework that safeguards fundamental rights without impeding technological innovation.[8]

3 METHODOLOGY

This study aims to analyse the legal framework governing criminal liability for generative artificial intelligence systems, focusing on the ChatGPT model as a contemporary practical example that reflects the legal and technical challenges in this field. To achieve this, the study employs a descriptive-analytical approach that examines national and international legislative texts and relevant legal principles, and reviews contemporary jurisprudential and legal literature on the nature of criminal liability amid the rapid development of AI technologies. Additionally, the study employs a comparative legal method, which systematically identifies and investigates specific areas where legal systems diverge and converge. The comparative criteria include factors such as the degree of criminal liability, regulatory protection, enforcement tactics, and ethical accountability standards. Using this framework, the study contrasts advanced European legislation, particularly the Artificial

5 ibid 9.

6 Omena Akpobome, 'The Impact of Emerging Technologies on Legal Frameworks: A Model for Adaptive Regulation' (2024) 5(7) International Journal of Research Publication and Reviews 5049. doi:10.55248/gengpi.5.1024.3012.

7 M Navaneeth, 'The Need for A Global Regulatory Framework for Artificial Intelligence: Implications of the European Union European Union Artificial Intelligence Act 2024' (Master's thesis, National University of Advanced Legal Studies 2024) 62-77; Mohammed Salem Alneyadi and others, 'The Crime of Electronic Blackmail in the Emirati Law' (2022 International Arab Conference on Information Technology (ACIT), Abu Dhabi, UAE, 22-24 November 2022). doi:10.1109/ACIT57182.2022.9994165.

8 Jennifer Kuzma and others, 'An Integrated Approach to Oversight Assessment for Emerging Technologies' in Gary E Marchant and Wendell Wallach (eds), *Emerging Technologies: Ethics, Law and Governance* (Routledge 2020) 1199. doi:10.1111/j.1539-6924.2008.01086.x.

Intelligence Act (AI Act),[9] with local legislative systems that are still in their infancy in addressing emerging technological challenges.[10]

The study, employing a critical approach, assesses the effectiveness of current legal frameworks in addressing AI-related harms, emphasising the challenges of establishing traditional causal links between actions and outcomes in algorithmic contexts. It explores the potential adoption of new liability models centred on preventive liability and criminal negligence, using case studies to illustrate how laws can evolve to balance societal protection with the promotion of technological innovation.

4 CRIMINAL LIABILITY DEFINITION FOR ARTIFICIAL MINDS AND CHATGPT

4.1. A Mind Without a Body: Who Prosecutes ChatGPT?

The phenomenon of ChatGPT vividly exemplifies the profound complexities artificial intelligence introduces into the criminal legal system.[11] This advanced linguistic system does not merely process data, but generates textual decisions that interact with humans and influence their cognitive, social, and legal realities.[12] When the generated text becomes capable of shaping convictions or guiding decisions, we are no longer dealing with a mere silent technical tool but a virtual mind without a body—one that redefines legal agency.[13] This raises fundamental questions about the nature and legal classification of artificial intelligence, especially in the absence of a unified definition within legal systems.[14] Traditional criminal models of intent, perpetrator identification, and liability must be reevaluated as AI's legal identity straddles the line between a human-controlled tool and an autonomous decision-maker.[15]

9 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 'Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828' (Artificial Intelligence Act - IA Act) [2024] OJ L 1689 <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> accessed 10 August 2025.

10 Navaneeth (n 7) 62-77.

11 Christiaan Mineur, 'Autonomous AI Technology and the Evolution of Legal Personhood in Criminal Law' (Master's thesis, University College Tilburg 2024) 4.

12 Iman M Al-Uqdah, 'Criminal Liability for Artificial Intelligence Application Crimes' (2025) 153 Journal of Law and Jurisprudence 39.

13 Mineur (n 11) 8.

14 Maxi Scherer, 'Artificial Intelligence and Legal Decision-Making: The Wide Open?' (2019) 36(5) Journal of international arbitration 541. doi:10.54648/joia2019028.

15 Jacob Turner, 'Legal Personality for AI' in Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Springer 2018) 175. doi:10.1007/978-3-319-96235-1_5.

When we ask, *"Who prosecutes ChatGPT?"*, we pierce through the veil of classical law and enter an unprecedented legal space.[16] The European Commission's 2021 proposal,[17] despite its effort to define AI broadly as systems capable of "generating outputs that affect the environment," fails to address the core dilemma: how to distinguish between basic AI and those complex generative systems that produce socially and legally impactful texts—like ChatGPT.[18] The challenge lies in three central characteristics: autonomy, interactivity, and opacity.[19] The term "autonomy" describes AI's capacity to act without direct oversight, not its intention. While opacity reflects the "black box" nature of its unpredictable algorithms, interactivity arises from daily user engagement.[20] These traits complicate the attribution of criminal liability, as the lines blur between programming error and human intention, between spontaneous output and directed decision.[21]

The dilemma posed by ChatGPT goes beyond legal debate into deep philosophical and ethical territory, shaking the foundations of traditional criminal concepts.[22] The absence of premeditation, the unpredictability of outputs, and the difficulty in identifying a clear actor, be it the developer, the user, or the owning company, reshapes the question of criminal liability.[23] It transforms it from a simple binary framework into a multi-layer network.[24] The traditional legal system, built on the formula "actor–victim–harm," is no longer adequate to encompass intelligent entities that commit crimes not in conventional ways but through knowledge flows open to interpretation.[25]

16 Amirreza Ahkami, 'AI and The European Union's Approach to Data Protection: The Case of Chat GPT' (Master's thesis, University of Padova 2024) 33-4.

17 European Commission, 'Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts' (COM/2021/206 final, 21 April 2021) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206> accessed 10 August 2025.

18 Navaneeth (n 7) 62-77.

19 Bram Vaassen, 'AI, Opacity, and Personal Autonomy' (2022) 35(4) Philosophy & Technology 89. doi:10.1007/s13347-022-00577-5.

20 Youliang Yuan and others, 'Does ChatGPT Know that it Does Not Know? Evaluating The Black-Box Calibration of Chatgpt' (2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), Torino, Italia, 20-25 May 2024) 5191. See also, Jingyu Wang and others, 'Network Meets ChatGPT: Intent Autonomous Management, Control and Operation' (2023) 8(3) Journal of Communications and Information Networks 340. doi:10.23919/JCIN.2023.10272352.

21 Briony Blackmore, 'Looking Beyond Blame and Praise: Analyzing Moral Responsibility in the Development and Deployment of AI Systems' (PhD thesis, University of Otago 2023) 12.

22 Al-Uqdah (n 12) 56.

23 Scherer (n 14) 542.

24 Xin Chen, 'Research on the Application of Intelligent ChatGPT in Computer Intelligent Computing System' (2023 IEEE 3rd International Conference on Data Science and Computer Application (ICDSCA), Dalian, China, 27-29 October 2023) 985. doi:10.1109/icdsca59871.2023.10392475.

25 Turner (n 15) 174.

Therefore, there is a pressing need for a new legal model based on "distributed shared accountability" and transparent oversight mechanisms that allow for tracing algorithmic decisions and assessing their legality and impact.[26] Ultimately, prosecuting a "mind without a body" is not a metaphor—it is a tangible necessity that demands an epistemological revolution in our understanding of law and a careful balance between technological innovation and legal protection of rights and freedoms in the digital age.[27]

4.2. When ChatGPT Speaks Without Criminal Intent: Can It Be Prosecuted?

We are not here to explore the intricate technical workings that give ChatGPT its structure, as these have been thoroughly explained by specialists and go beyond the concerns of criminal jurisprudence.[28] Even in cases where there is no criminal intent or will, what matters is how AI-generated linguistic outputs affect the legal system.[29] The machine learning model does not operate on explicit logical rules; rather, it follows a probabilistic inductive approach, extracting linguistic patterns from billions of examples without "understanding" them. ChatGPT neither knows the truth nor intends to lie, yet it can generate harmful, misleading, or inflammatory content.[30] AI thus generates "speech without reason" and "action without intent," undermining traditional legal theories that attribute criminal liability exclusively to human consciousness and rational awareness.[31]

When ChatGPT generates illegal or criminally consequential content, the challenge of identifying the responsible actor emerges. Is it the model itself? The developers? The owning company? Or the users? Like a "blind painter," the model uses probabilistic estimates that change as its inputs change to create linguistic portraits without knowing why or for what purpose.[32] To maximise flexibility and generative capacity rather than the logical consistency required for legal liability, its architecture is deliberately opaque.[33] Under the new legal concept of distributed liability, traditional actors cannot be held

26 Kuzma and others (n 8) 1198.

27 Akpobome (n 6) 5050.

28 Kalliopi Terzidou, 'Generative AI for the Legal Profession: Facing the Implications of the Use of ChatGPT Through an Intradisciplinary Approach' (*Media Laws*, 8 September 2023) 4 <https://www.medialaws.eu/generative-ai-for-the-legal-profession-facing-the-implications-of-the-use-of-chatgpt-through-an-intradisciplinary-approach/> accessed 10 August 2025.

29 Aslihan Asil and Thomas G Wollmann, 'Can Machines Commit Crimes Under US Antitrust Laws?' (2024) 3(1) The University of Chicago Business Law Review 6 <https://businesslawreview.uchicago.edu/print-archive/can-machines-commit-crimes-under-us-antitrust-laws > accessed 10 August 2025.

30 Terzidou (n 28) 2.

31 Lawrence B Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 North Carolina Law Review 1272-3.

32 Sarah Muller, 'Visual Silence in the Language Portrait: Analyzing young People's Representations of their Linguistic Repertoires' (2022) 25(10) International Journal of Bilingual Education and Bilingualism 3646. doi:10.1080/13670050.2022.2072170.

33 Blackmore (n 21) 44.

accountable; instead, accountability must be traced across multiple stakeholders, from user interfaces to training environments.[34]

Even though ChatGPT does not aim to break the law or incite, its reliance on linguistic patterns can yield results with significant ethical or legal implications. The legal challenge lies here: how do we assign accountability to a system that lacks will, yet produces effects? The answer requires moving beyond traditional liability frameworks by developing new rules that hold developers and operators accountable and enforce proactive oversight of inputs and algorithms. The legal focus must shift from "the actor's intent" to "design and operational responsibility," and from the "criminal mind" to a system of "digital governance." This model does not err because it chooses to, but because it lacks the capacity to distinguish right from wrong—necessitating legislation that redefines the relationship between technology and accountability under a new logic.[35]

ChatGPT stands at the threshold of artificial consciousness, in a legal grey zone that criminal justice systems are not yet prepared to handle. It does not think or comprehend, but it generates discourse that simulates thought. It is neither a traditional actor nor a mere tool—it is a linguistic entity that challenges settled legal classifications. In this sense, artificial intelligence functions more as a mirror exposing the inadequacies of our laws than as a standalone problem. The challenge lies not in how "intelligent" it is, but in how legally "prepared" we are to incorporate it into our network of criminal concepts. A reevaluation of crime and punishment in which actors may be nonhuman, and liability arises from error, probability, or unanticipated consequences rather than conscious intent, would result from failing to act, risking the prosecution of algorithms for unintended outputs.

4.3. Attributing Criminal Liability in the Age of Intelligent Machines

Artificial intelligence systems pose a genuine challenge to the traditional criminal liability framework, which is built upon the pillars of actus reus (the act), mens rea (intent), and will.[36] These intelligent entities are neither human nor self-aware nor criminally intent; rather, they are digital tools that generate unpredictable behaviours that are difficult to foresee accurately.[37] As reliance on these complex systems—operating on probabilistic rather than explicit logical bases—increases, a central legal question arises: How can liability for harm caused by these systems be assigned when humans lack full control over their behaviour? ChatGPT and similar language models do not rely on true understanding. Still, on intricate statistical patterns they neither comprehend nor can

34 Dirk A Zetzsche, Ross P Buckley and Douglas W Arner, 'The Distributed Liability of Distributed Ledgers: Legal Risks of Blockchain' (2018) 4 University of Illinois Law Review 1386. doi:10.2139/ssrn.3018214.

35 Akpobome (n 6) 5050.

36 Mineur (n 11) 22.

37 Asil and Wollmann (n 29) 20.

explain, making prosecution for harmful actions impossible under traditional criminal law concepts.[38] Hence, our legal system faces a profound crisis in how to classify these non-living entities that produce real-world effects and in defining a legal framework balancing technological progress with justice.[39]

Confronting this dilemma, international approaches, especially within the European Union, have emphasised the importance of strict civil liability as a practical mechanism to protect victims.[40] Directive 85/374/EEC on defective products is a law that holds manufacturers liable for damages without needing proof of fault or intent.[41] Traditional concepts of "defect" and "causality" are blurred by the complexity of AI since intelligent systems are dynamic entities with probabilistic behaviours rather than traditional products.[42] This prompted the European Commission in 2022 to propose comprehensive legal updates addressing "smart products".[43] These updates aim to broaden legal protection and shift the burden of proof onto producers by presuming a link between defect and damage automatically, and imposing economic responsibility on producers for risks associated with these systems—even when their behaviour is unpredictable or unforeseeable.[44] This reflects a fundamental shift in liability philosophy—from focusing on the actor's intent to ensuring effective compensation for victims regardless of the actor's awareness.[45]

Since AI is an unconscious entity lacking intent or will, the law must transcend traditional concepts grounded in these elements and develop a hybrid legal framework combining strict civil responsibility, regulatory liability, and supervisory oversight.[46] Producers or developers oversee the implementation of safety precautions and ensure clear standards, aiming to prevent harm and provide victims with straightforward compensation.[47] Instead of prosecuting algorithms without understanding how they work, the AI approach

38 Mineur (n 11) 19.
39 Scherer (n 14) 542.
40 Ahkami (n 16) 39.
41 Council Directive 85/374/EEC of 25 July 1985 'On the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products' [1985] OJ L 210/29; Fidelma White, 'Directive 85/374/EEC Concerning Liability for Defective Products: In the Name of Harmonisation, the Internal Market and Consumer Protection' in Paula Giliker (ed), *Research Handbook on EU Tort Law* (Edward Elgar 2017) 128. doi:10.4337/9781785365720.00013.
42 Keith Darlington, 'Aspects of Intelligent Systems Explanation' (2013) 1(2) Universal Journal of Control and Automation 47. doi:10.13189/ujca.2013.010204.
43 European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on Horizontal Cybersecurity Requirements for Products with Digital Elements and Amending Regulation (EU) 2019/1020' (COM/2022/454 final, 15 September 2022) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52022PC0454> accessed 10 August 2025; Ahkami (n 16) 39.
44 Navaneeth (7) 62-77.
45 Cheatham, Javanmardian and Samandari (n 4) 9.
46 Kuzma and others (n 8) 1197.
47 Oscar Oviedo-Trespalacios and others, 'The Risks of Using ChatGPT to Obtain Common Safety-Related Information and Advice' (2023) 167 Safety Science 106244. doi:10.1016/j.ssci.2023.106244.

emphasises accountability for system design and operation, prevention, and compensation.[48] Without renewing our legal frameworks, we risk a future in which algorithms are prosecuted for unintended or unforeseen actions, opening a dangerous legal vacuum that threatens the application of justice.[49]

5 CHALLENGES IN ASSESSING THE CRIMINAL LIABILITY OF ARTIFICIAL INTELLIGENCE

Conventional criminal law faces unprecedented challenges because of the rise of ChatGPT and other generative AI systems. At the core of this disruption is the causality dilemma: it is impossible to use conventional legal methods to link AI actions to outcomes. The fact that AI operates without conscious intent, which contradicts conventional notions of guilt, exacerbates the *mens rea* dilemma. When taken as a whole, these crises highlight the pressing need to reconsider liability assessment in the age of intelligent machines.

5.1. The Crisis of Criminal Law in the Era of ChatGPT

The digital revolution is causing a major crisis for criminal law, which is still confined to traditional frameworks based on human consciousness, intent, and choice.[50] The criminal system is governed by well-established principles such as legality, personal blame, and the presumption of innocence, designed to address human actions with clear motives. Still, it is ill-equipped to accommodate acts generated by generative AI systems like ChatGPT.[51] Determining criminal liability in systems with self-learning capabilities and programmers' autonomy is challenging because of unpredictable behaviours that cannot be traced to specific human actors.[52] This clash represents the "shock of modernity" described by the Italian thinker Federico Stella, in which traditional criminal concepts such as intent and awareness lose their relevance when confronted with technology characterised by behavioural ambiguity and unpredictability.[53]

AI-induced harm challenges the evidentiary and proof mechanisms of criminal law because algorithmic biases or training errors make it more difficult to assign blame and to establish a

48 Sonia K Katyal, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66(1) UCLA Law Review 90.

49 Rebecca Crootof, '"Cyborg Justice" and the Risk of Technological-Legal Lock-In' (2019) 119(7) Columbia Law Review 249.

50 Sergio Carrera, Valsamis Mitsilegas and Marco Stefan, *Criminal Justice, Fundamental Rights and the Rule of Law in the Digital Age: Report of a CEPS and QMUL Task Force* (CEPS 2021). 62.

51 Blackmore (n 21) 44.

52 Monika Simmler, 'Responsibility Gap or Responsibility Shift? The Attribution of Criminal Responsibility in Human-Machine Interaction' (2024) 27(6) Information, Communication & Society 1145. doi:10.1080/1369118X.2023.2239895.

53 Solum (n 31) 1274.

direct causal relationship between the action and the harm.[54] In a participatory, multi-agent technological environment, personal liability becomes unclear as multiple parties' acts and omissions intertwine, weakening courts' certainty and threatening the principle of predictability of outcomes.[55] For instance, output generated by ChatGPT may be used to incite, defame, or commit fraud, but the absence of direct human intent complicates criminal accountability and exposes the limitations of current legal tools to address these complexities.[56]

The natural outcome of this crisis is a clear and troubling legal vacuum, where current criminal law lacks explicit and comprehensive rules addressing acts resulting from AI that lack awareness or will.[57] These vacuum places justice in a predicament: AI itself cannot bear criminal liability, nor can the traditional responsible human—whether developer or user—be easily held liable due to algorithmic complexity and the difficulty of proving fault and intent.[58] Consequently, today's digital reality demands a rethinking of the very definitions of crime and the principles of liability, opening the door to profound legal and philosophical debates about the limits of criminal law and how to develop a regulatory framework that balances societal protection from AI risks with encouraging innovation without threatening legal security and justice.[59]

5.2. The Causality Dilemma in the Era of Generative Artificial Intelligence

Generative AI systems like ChatGPT challenge traditional legal notions of causality.[60] Criminal law's deterministic model linking human acts to outcomes struggles against AI's probabilistic algorithms,[61] which produce unexpected outcomes without a human actor,[62] undermining classical proof and responsibility frameworks.[63]

54 Henrique Manuel Gil Martins, 'Liability Implications of Artificial Intelligence use in Health: Fault and Risk in Public Sector Healthcare' (Master's thesis, Universidade Catolica Portuguesa 2020) 31-2; Sander Beckers, Hana Chockler and Joseph Halpern, 'A Causal Analysis of Harm' (Advances in Neural Information Processing Systems 35: 36th Conference on Neural Information Processing Systems (NeurIPS 2022), 28 November - 9 December 2022, New Orleans, Louisiana, USA) 2368.

55 Oren Griffiths and Anna Thorwart, 'Effects of Outcome Predictability on Human Learning' (2017) 8 Frontiers in Psychology 514. doi:10.3389/fpsyg.2017.00511.

56 Akpobome (n 6) 5051.

57 Solum (n 31) 1273-4; Zetzsche, Buckley and Arner (n 34) 1286.

58 Nora Osmani, 'The Complexity of Criminal Liability of AI Systems' (2020) 14(1) Masaryk University Journal of Law and Technology 59. doi:10.5817/mujlt2020-1-3.

59 Cheatham, Javanmardian and Samandari (n 4) 8.

60 Blackmore (n 21) 45.

61 Emad H Atiq, 'How Folk Beliefs about Free Will Influence Sentencing: A New Target for the Neuro-Determinist Critics of Criminal Law' (2013) 16(3) New Criminal Law Review 449. doi:10.1525/nclr.2013.16.3.449.

62 Douglas C Youvan, 'Reconciling the Probabilistic and Deterministic: Exploring Complexity, Emergence, and Uncertainty in Nature, AI, and Human Cognition' (Research Gate, October 2024) 6. doi:10.13140/RG.2.2.16020.10880.

63 ibid 7-8.

It is challenging to prove a causal link between the actions of generative AI systems and the outcomes they generate due to their technical complexity. This puts conventional legal reasoning to the test.[64] In the past, courts used instruments like digital "black boxes" that capture event data to piece together the sequence of an incident and identify its cause.[65] The algorithm's internal decision-making processes, however, are based on complex, nonlinear probabilistic models, and these tools only provide preliminary indications.[66] Experts and judges are unable to conclusively determine whether an error with ChatGPT results from a programming error, bias in the training data, or user behaviour.[67] Such ambiguities make it more difficult to establish clear causation, particularly in cases of harmful content generation, and complicate the application of conventional principles that rely on a clear connection between action and result.[68]

Criminal law faces an epistemic crisis due to this causality conundrum, compelling a reexamination of its central concepts of intent, causation, and liability.[69] The nomological-deductive model that judicial systems employ excludes other possible explanations and necessitates precise scientific law and a logical causal relationship between an action and its result.[70] Generative AI undermines this model, as even system developers cannot precisely identify the causes of the outcomes they produce, rendering actions into multidimensional probabilistic outcomes.[71] Legal systems must accept "flexible causality", place blame on programmers, users, and AI systems, and update evidentiary standards to account for social and technical context to ensure criminal justice.[72] Law must adapt to this new reality, avoiding confinement within rigid models that fail to accommodate the algorithmic revolution, thus preserving the essence of justice and individual rights in the age of artificial intelligence.[73]

64 Beckers, Chockler and Halpern (n 54) 2368.
65 Yuan and others (n 20) 5192.
66 Wang-Ji Yan and others, 'Navigating Uncertainties in Machine Learning for Structural Dynamics: A Comprehensive Review of Probabilistic and Non-Probabilistic Approaches in Forward and Inverse Problems' (arXiv:2408.08629, 16 August 2024) 16. doi:10.48550/arXiv.2408.08629.
67 Youvan (n 62) 6.
68 Osmani (n 58) 63.
69 Blackmore (n 21) 45.
70 Rocco Neri, 'Judging Beyond any Reasonable Doubt: A Logic and Epistemological Rule' (2024) 7 Quaestio Facti: Revista internacional sobre razonamiento probatorio 49-50. doi:10.33115/udg_bib/qf.i7.23028.
71 Atiq (n 61) 449.
72 Osmani (n 58) 63.
73 Beckers, Chockler and Halpern (n 54) 2371.

5.3. ChatGPT and the Mens Rea Dilemma

Amid the rapid development of generative artificial intelligence, the ChatGPT model has emerged as an icon of digital transformation in natural language processing.[74] However, this model embodies a fundamental legal crisis that touches the core of causal proof—a foundational element in determining criminal liability. Traditional legal concepts rely on a clear and direct causal relationship between a human act and a harmful result, based on a nomological model that infers harm from a specific act carried out with awareness and intent.[75] By contrast, ChatGPT, as a complex algorithmic system operating through millions of probabilities, does not produce actions with intent or purpose. Its outputs are based on data and deep learning, not on conscious decisions.[76]

For example, if ChatGPT generates false or inciting content that causes moral or material damage, a complex question arises: can it be proven that the model's output directly caused these outcomes? Or is the human user who deployed the content responsible? In such cases, traditional legal rules requiring the attribution of harm to a conscious being become difficult to apply when faced with the opacity of an "artificial mind" that possesses neither will nor intention.[77]

The mental element (*mens rea*) is a cornerstone of criminal liability, typically manifested in intent or negligence.[78] However, when discussing ChatGPT, these concepts begin to dissolve, as the model has no will or consciousness and thus cannot possess intent like a human actor.[79] This raises the question of whether producers or developers are exonerated. Not necessarily, as existing doctrines provide alternative bases for attribution, the law might invoke the notion of *probable intent*, holding developers or users accountable if they could reasonably foresee the harm as a potential consequence of using the model—even in the absence of malicious intent.[80]

Suppose a developer or company releases a version of ChatGPT capable of generating inciting or offensive content, knowing that it could be misused, yet fails to implement sufficient preventive measures. Here, *probable intent* is realised, as they are expected to foresee the risk and prevent it.[81] Conversely, if harm results from unexpected use within a

74 Muller (n 32) 22.
75 Solum (n 31) 1274.
76 Mineur (n 11) 22.
77 Vaassen (n 19) 89-90.
78 Mineur (n 11) 22.
79 Soraj Hongladarom and Auriane Van der Vaeren, 'ChatGPT, Postphenomenology, and the Human-Technology-Nature Relations' (2024) 2 Journal of Human-Technology Relations 9. doi:10.59490/jhtr.2024.2.7386.
80 ibid 11-2.
81 Gabriel Hallevy, 'The Criminal Liability of Artificial Intelligence Entities-From Science Fiction to Legal Social Control' (2010) 4(2) Akron Intellectual Property Journal 191.

wide range of possible outcomes, assigning liability becomes more complex.[82] The model's opacity adds to the challenge, as even the developers might not be able to explain why certain content was generated, making it difficult to determine whether the act constituted criminal negligence or an acceptable technical risk.[83]

To address these complexities, the law must go beyond mere adjustments to evidentiary tools. It must construct a new regulatory system that considers the nature of generative artificial intelligence.[84] For instance, clear legal frameworks should be established to govern the development and operation of models like ChatGPT, imposing strict standards for digital safety and governance responsibility, and defining the limits of acceptable risks—or what might be termed the "permissible risk zone".[85]

Imagine a company developing ChatGPT that implements preventive measures such as content filtering and usage management to prevent harmful outputs. These measures would delineate the permissible risk zone, serving as the dividing line between acceptable technical error and liability overreach.[86] Liability would arise only when these boundaries are breached, due to a lack of precautionary measures or willful disregard for potential risks.[87]

Thus, the notions of intent and negligence must be reformulated within a broader framework that accounts for varying degrees of foreseeability and probability, while recognising the unique nature of "actions" generated by a non-conscious mathematical system.[88] The law now faces a fundamental philosophical challenge—not merely redrawing evidentiary tools but reconstructing its foundational perceptions of criminal liability in an era where artificial intelligence has become an independent, complex actor generating actions and consequences that transcend traditional concepts of will and intent.[89]

6 THE LEGAL AND REGULATORY CHALLENGES OF GENERATIVE AI

As generative artificial intelligence, like ChatGPT, grows quickly, new legal issues regarding legal risks and responsibility for generated content surface. Determining criminal liability is challenging due to a regulatory gap left by current frameworks that have not kept up with the rate of innovation. To suggest practical ways to ensure the safe and lawful use of this

82 ibid 194.

83 Yuntao Wang and others, 'A Survey on ChatGPT: AI-Generated Contents, Challenges, and Solutions' (2023) 4 IEEE Open Journal of the Computer Society 297. doi:10.48550/arXiv.2305.18339.

84 Xukang Wang and Ying Cheng Wu, 'Balancing Innovation and Regulation in the Age of Generative Artificial Intelligence' (2024) 14 Journal of Information Policy 397-8. doi:10.5325/jinfopoli.14.2024.0012.

85 ibid 394.

86 Wang and others (n 83) 280.

87 Wang and Wu (n 84) 394-5.

88 Mineur (n 11) 22.

89 Hallevy (n 81) 186-7.

technology, this theme investigates the nature of algorithmic intelligence, evaluates legal risks, and looks at regulatory gaps.

6.1. The Algorithmic Regulatory Gap and Legal Accountability of ChatGPT

Generative AI models, led by ChatGPT, are characterised by a unique cognitive nature, as they operate on complex algorithms that form what is known as the "black box," whose inner workings are difficult to explore or whose outputs are hard to interpret precisely.[90] In the absence of a comprehensive scientific framework clarifying the causal pathways these models follow in making decisions, users face a fundamental challenge: the inability to understand how the model selects certain phrases or information over others, or to interpret the logic underlying the arrangement of elements in the response.[91] For example, when ChatGPT provides a legal recommendation in a criminal case, neither the judge nor the lawyer has access to the detailed reasons that led the model to formulate this recommendation, which obstructs the possibility of accountability if an error or harm occurs.[92] In this sense, this interpretive gap undermines trust in AI outputs and poses fundamental challenges in terms of transparency and legal accountability.[93]

A significant portion of errors produced by models like ChatGPT stems from intertwined technical and legal factors.[94] On the technical level, some errors relate to the design of algorithmic architecture, where certain arrangements in the neural network layers can lead to unintended implicit biases, such as the model discriminating between individuals based on gender or race.[95] At the data level, AI models rely on massive amounts of text drawn from diverse online sources, which may carry cultural or ethical biases or contain inaccurate legal information, thereby transferring these biases into the model's outputs.[96]

Given these multiple challenges, there is an urgent need to develop a strict legal regulatory framework governing the operation of generative AI models and ensures the safety and reliability of their outputs, especially in highly sensitive contexts such as the legal field.[97] The European AI Act represents a pioneering step in this direction,[98] imposing on developers of models like OpenAI a comprehensive risk management obligation, requiring them to ensure training data quality, prepare technical documentation that clarifies decision-

90 Yuan and others (n 20) 5192.
91 Muller (n 32) 3646.
92 ibid 3647.
93 Simmler (n 52) 1146.
94 Amos Azaria, Rina Azoulay and Shulamit Reches, 'ChatGPT is a Remarkable Tool—for Experts' (2024) 6(1) Data Intelligence 241. doi:10.1162/dint_a_00235.
95 Gabbrielle M Johnson, 'Algorithmic Bias: On the Implicit Biases of Social Technology' (2021) 198(10) Synthese 9947. doi:10.1007/s11229-020-02696-y.
96 ibid 9952.
97 Navaneeth (n 7) 62-77.
98 Regulation (EU) 2024/1689 (n 9).

making logic, maintain accurate logs of system outputs, and mandating human oversight of system outputs in high-risk cases.[99] Legal requirements for risk reduction or elimination could make the judiciary less accountable, calling for new models of legal accountability that account for the cognitive differences between humans and machines.[100] This necessitates establishing mechanisms for interpretation and analysis that assist users in consciously and thoughtfully understanding and evaluating the model's results.

6.2. Algorithmic Intelligence and Legitimate Risks

The development of generative artificial intelligence systems, such as the ChatGPT platform, raises novel legal challenges concerning liability for damages resulting from their use.[101] Even if the manufacturing company complies fully with all prescribed technical and regulatory standards, the pressing question remains: Is such compliance sufficient to absolve it of legal liability? Legal and philosophical experience indicates that formal adherence to rules does not necessarily prevent harm, especially when algorithms are granted quasi-autonomous power to make complex decisions, as is the case with these systems operating within unpredictable environments and multifaceted causes.[102] For example, an algorithmic error in interpreting a query or generating inaccurate content may cause serious psychological or social harm that cannot be fully anticipated, thereby fueling debate over whether technical standards alone provide an adequate legal defence.[103]

In legal theory, compliance with rules is recognised as a necessary but insufficient condition for relieving liability, as courts also rely on the "reasonable actor" or "person of similar position and competence" standard.[104] This requires assessing whether the manufacturer took the necessary professional and prudent measures to avoid harm. In complex technical industries, jurisprudence acknowledges the existence of "residual risks" or "acceptable risks"—those that cannot be eliminated except at exorbitant costs or at the expense of technological progress.[105] According to Amy Stein,[106] these risks represent an implicitly accepted zone of risky behaviours, provided they remain within the bounds of control and reasonable precautions, reflecting the reality of technological development and balancing innovation with safety.

99 Kuzma and others (n 8) 1202.
100 Akpobome (n 6) 5051.
101 Johnson (n 95) 9962.
102 Simmler (n 52) 1145.
103 Mike Ananny, 'Seeing like an Algorithmic Error: What Are Algorithmic Mistakes, Why Do They Matter, How Might They Be Public Problem?' (2022) 24(Spec) Yale Journal of Law & Technology 48.
104 ibid 52-3.
105 Azaria, Azoulay and Reches (n 94) 241.
106 Amy L Stein, 'Assuming the Risks of Artificial Intelligence' (2022) 102 Boston University Law Review 983.

On the legislative front, the European Artificial Intelligence Act (AI Act)[107] exemplifies legal evolution that abandons the ideal of eliminating risks and instead adopts a pragmatic philosophy aimed at minimising risks as much as possible while acknowledging their persistent possibility.[108] This law not only imposes strict technical standards but also emphasises algorithmic transparency, continuous system performance monitoring, and the interpretability of decision-making mechanisms.[109] This places a broader duty on manufacturers, extending beyond formal compliance to proactive professional conduct and adaptation to scientific developments.[110] For instance, if damage results from a fault in an algorithm trained on unbalanced data, the manufacturer bears the burden of proving that it took all necessary measures to prevent such errors or mitigate their effects.

The fundamental dilemma lies in delineating the boundary between harm accepted as part of "residual risks" and unjustified failures in design or operation.[111] Accordingly, proposals have emerged to assess liability based on comparing overall system performance against a "safe model" standard; if harmful error rates exceed a certain threshold, the design is considered defective, even if individual errors may be technically justified.[112] However, this approach faces technical challenges related to real-time monitoring, as well as political and ethical considerations involving acceptance of any potential harm to human life.[113] Considering this, there is a pressing need to develop an integrated legal framework that connects technical compliance, professional responsibility, and ethical caution — affirming that justice cannot rest solely on textual adherence but requires a deeper perspective aligned with the complexities and rapid evolution of generative AI.

7 THE FUTURE OF CRIMINAL LIABILITY FOR ALGORITHMIC HARM

Traditional ideas of criminal liability are being called into question by the emergence of generative artificial intelligence, particularly as it permeates business operations. It is becoming more difficult to assign blame for damage brought on by autonomous AI decisions. To predict the future of criminal law amid profound technological change, emerging research aims to reconstruct legal frameworks suited to corporate AI.

107 Regulation (EU) 2024/1689 (n 9).
108 Navaneeth (n 7) 62-77.
109 ibid
110 ibid
111 Stein (n 106) 988.
112 ibid
113 Simmler (n 52) 1145.

7.1. Rebuilding Generative AI's Criminal Liability

Despite notable progress in developing frameworks for criminal liability in the age of artificial intelligence, the legal pathway to holding producers of AI systems, such as ChatGPT, criminally responsible remains complex and ambiguous. Features of these systems—from unpredictability to the opacity of algorithmic decision-making.[114] This creates an epistemic gap between human action and harmful outcome, undermining the establishment of a clear causal link and weakening the logic of blame based on negligence or fault.[115] This phenomenon has been described in legal literature as the "control gap crisis," which criminal law faces in a technological risk society, where traditional models centred on criminal conduct and free will fail to accommodate the emerging complexities of algorithmic actions. For instance, a decision by an autonomous vehicle's algorithm may cause an accident without direct human intervention, raising profound questions about criminal liability.

Traditional criminal liability, which is predicated on intent or negligence, struggles to handle algorithmic decisions. Modern approaches emphasise duty of care violations when AI systems endanger fundamental rights, such as psychological or physical safety, placing a higher priority on preventive liability. Laws such as Article 12 of the UAE consumer protection law No (15) of 2020, which penalise dangerous products but still partially exempts software-only AI, reflect this trend. Thus, liability shifts from fault and intent to adherence to preventive and precautionary obligations.

The particularity of algorithmic harm calls for legislative renewal to keep pace with technological changes by treating AI as a legal entity with tangible effects—even absent physical embodiment. The European AI Act represents a significant step forward, imposing effective, balanced, and deterrent sanctions on violators of AI system regulations.[116] The concept of a "legal warning" model also emerges, criminalising behaviours such as neglecting security measures or failing to update systems, with a combined regime of administrative oversight and criminal penalties to manage risks.[117] Within this framework, authorities may be empowered to issue legal orders mandating additional testing, compulsory updates, or partial system suspension, with violations triggering criminal liability—thereby strengthening societal legal protection.[118]

The nature of algorithmic harm demands establishing a novel legal concept that redefines the relationship between algorithmic acts and legal efficacy, ensuring harm does not escape accountability under the guise of technology and innovation.[119] Accordingly,

114 Vaassen (n 19) 89-90.
115 Simmler (n 52) 1146.
116 Ahkami (n 16) 39.
117 Kuzma and others (n 8) 1201.
118 Cheatham, Javanmardian and Samandari (n 4) 8.
119 Katyal (n 48) 91-2.

enhancing algorithmic transparency by obliging AI producers, such as ChatGPT, to regularly disclose potential risks supports a "democratisation of risk management" approach.[120] This recognises the right of society and legislators to be informed of the potential impacts of technology permeating all aspects of life. Consequently, it becomes essential to combine principles of strict or risk-based liability with advanced preventive legal mechanisms to balance the promotion of innovation with the protection of fundamental legal values such as safety and dignity, thereby achieving legal stability and criminal justice in addressing AI-caused harms.

7.2. Future Studies: Corporate Criminal Law and AI Liability

Traditional ideas of criminal liability are coming under increasing pressure from the emergence of generative artificial intelligence, especially as it becomes increasingly ingrained in business settings.[121] The complexity of determining who is responsible for damage brought about by autonomous AI decisions has led to new research aimed at rebuilding legal frameworks suitable for corporate AI.[122] This changing course aims to anticipate the future of criminal law and guarantee its flexibility in response to the significant technological advancements influencing contemporary responsibility and governance.[123]

Future studies will focus on identifying the parties accountable for damage brought about by generative AI systems in commercial settings.[124] It is expected that questions will arise about the extent to which traditional notions of corporate criminal liability, such as vicarious responsibility or failure to supervise, can address harms caused solely by AI algorithms. It might be necessary to develop new legal frameworks that enable courts to hold companies responsible for algorithmic decisions.[125]

Future research is expected to focus heavily on evaluating businesses' preventive responsibilities when using artificial intelligence. This could mean examining how much corporate governance, risk monitoring, and legal compliance are required of businesses, and how failing to comply could be interpreted as criminal negligence or complicity under corporate criminal law.[126]

Comparative analyses of various legal systems are also likely to become increasingly important, offering insights into how conventional legal doctrines interact with the novel

120 Cheatham, Javanmardian and Samandari (n 4) 8.
121 Mohammad Amin Alkrisheh, 'Criminal Protection of Corporate Websites: An Analytical Study' (2022) 11(3) Journal of Governance and Regulation 148. doi:10.22495/jgrv11i3art12.
122 Osmani (n 58) 59.
123 Kuzma and others (n 8) 1199.
124 Mineur (n 11) 22.
125 Hongladarom and Van der Vaeren (n 79) 11-2.
126 Stavros Kalogiannidis and others, 'The Role of Artificial Intelligence Technology in Predictive Risk Assessment for Business Continuity: A Case Study of Greece' (2024) 12(2) Risks 19. doi:10.3390/risks12020019.

challenges posed by AI-related harms. Such studies may support lawmakers and prosecutors in creating unified regulatory standards and future guidelines to control corporate risks related to artificial intelligence.

The notion of legal capacity for systems like ChatGPT involves treating AI as a potential legal entity with unique rights and responsibilities.[127] At present, global criminal law views AI as a tool for holding humans accountable, even as scholars point out that the existing legal system is not yet fully equipped to address accountability and regulatory questions raised by AI. An industrial or electronic form of legal capacity has been proposed as a potential model for structuring this recognition. Moreover, ChatGPT or AI systems would be punished by operational or technical limitations, such as halting operations or requiring updates, that are intended to prevent harm and ensure compliance, rather than the typical human-centred sanctions.

8 CONCLUSIONS AND RECOMMENDATIONS

The examination of criminal liability for artificial intelligence technologies, especially conversational systems like ChatGPT, reveals profoundly beneficial aspects that transcend the theoretical framework. Legislators are forced to reconsider traditional ideas of criminal liability as well as the *actus reus* (the act) and *mens rea* (intent) components of crime, considering the increasingly intertwined behaviour of humans and machines. It also provides legal professionals with new analytical tools to understand technological risks and assess their obligations in complex digital environments. Therefore, this study contributes to the development of future criminal policy toward a more flexible and equitable system in the age of artificial intelligence, while also improving scholarly discourse.

Fault-based liability is undermined by the difficulty of assigning criminal liability for algorithmic harm, as AI decisions often lack the traditional human intent required for criminal liability. Fundamental ideas in criminal law must be examined in light of the complexity and loss of control over technology. Legal standards that integrate technological tools to enhance preventive oversight and transparency are necessary to ensure accountability for harm caused by AI while balancing innovation and the defence of fundamental rights.

ChatGPT's legal capability could be electronically acknowledged, and operational or technical safeguards could ensure compliance and prevent harm, thereby ensuring accountability. Establishing a direct causal link between algorithmic actions and harmful outcomes is complicated by the opaque, complex nature of AI systems' decision-making processes. Traditional criminal law struggles to effectively address unpredictable, uncontrollable technological risks. There is a growing need to shift from fault-based

127 Yuan and others (n 20).

liability to a preventive liability model that emphasises the failure to implement adequate protective measures. Existing legal frameworks, including consumer protection laws, are inadequate for addressing the unique challenges posed by AI, especially for software that lacks physical integration. Transparency and regulatory oversight are critical to managing AI-related risks, with regular disclosure and mandatory updates serving as key tools to mitigate potential harm.

It is recommended that a legal framework be established that recognises AI's electronic capabilities and outlines operational or technical procedures as accountability mechanisms. Legal frameworks should be modernised to recognise AI as a distinct legal entity, capable of bearing responsibilities regardless of its physical form, with clear duties imposed on developers and users. A preventive liability approach should be adopted, criminalising negligence in failing to ensure safety measures and system updates, with deterrent penalties. AI developers must be required to provide regular transparency reports about potential risks, supporting broader societal and legislative oversight mechanisms. Administrative authorities should be empowered to monitor AI systems, enforce mandatory testing and updates, and impose partial shutdowns, when necessary, with criminal sanctions for non-compliance to ensure effective protection.

REFERENCES

1. Ahkami A, 'AI and the European Union's Approach to Data Protection: The Case of Chat GPT' (Master's thesis, University of Padova 2024)
2. Akpobome O, 'The Impact of Emerging Technologies on Legal Frameworks: A Model for Adaptive Regulation' (2024) 5(7) International Journal of Research Publication and Reviews 5046. doi:10.55248/gengpi.5.1024.3012
3. Alkrisheh MA, 'Criminal Protection of Corporate Websites: An Analytical Study' (2022) 11(3) Journal of Governance and Regulation 148. doi:10.22495/jgrv11i3art12
4. Alneyadi MS and others, 'The Crime of Electronic Blackmail in the Emirati Law' (2022 International Arab Conference on Information Technology (ACIT), Abu Dhabi, UAE, 22-24 November 2022). doi:10.1109/ACIT57182.2022.9994165
5. Al-Uqdah IM, 'Criminal Liability for Artificial Intelligence Application Crimes' (2025) 153 Journal of Law and Jurisprudence 39
6. Ananny M, 'Seeing like an Algorithmic Error: What Are Algorithmic Mistakes, Why Do They Matter, How Might They Be Public Problem?' (2022) 24(Spec) Yale Journal of Law & Technology 342
7. Asil A and Wollmann TG, 'Can Machines Commit Crimes Under US Antitrust Laws?' (2024) 3(1) The University of Chicago Business Law Review 1

8. Atiq EH, 'How Folk Beliefs about Free Will Influence Sentencing: A New Target for the Neuro-Determinist Critics of Criminal Law' (2013) 16(3) New Criminal Law Review 449. doi:10.1525/nclr.2013.16.3.449

9. Azaria A, Azoulay R and Reches S, "ChatGPT is a Remarkable Tool—for Experts' (2024) 6(1) Data Intelligence 240. doi:10.1162/dint_a_00235

10. Beckers S, Chockler H and Halpern J, 'A Causal Analysis of Harm' (Advances in Neural Information Processing Systems 35: 36th Conference on Neural Information Processing Systems (NeurIPS 2022), 28 November - 9 December 2022, New Orleans, Louisiana, USA) 2365

11. Biju PR and Gayathri O, 'Algorithmic Solutions, Subjectivity and Decision Errors: A Study of AI Accountability' (2025) 27(5) Digital Policy, Regulation and Governance 523. doi:10.1108/DPRG-05-2024-0090

12. Blackmore B, 'Looking Beyond Blame and Praise: Analysing Moral Responsibility in the Development and Deployment of AI Systems' (PhD thesis, University of Otago 2023)

13. Carrera S, Mitsilegas V and Stefan M, *Criminal Justice, Fundamental Rights and the Rule of Law in the Digital Age: Report of a CEPS and QMUL Task Force* (CEPS 2021)

14. Cheatham B, Javanmardian K and Samandari H, 'Confronting the Risks of Artificial Intelligence' (2019) 55(2) McKinsey Quarterly 38

15. Chen X, 'Research on the Application of Intelligent ChatGPT in Computer Intelligent Computing System' (2023 IEEE 3rd International Conference on Data Science and Computer Application (ICDSCA), 27-29 October 2023) 985. DOi:10.1109/icdsca59871.2023.10392475

16. Crootof R, '"Cyborg Justice" and the Risk of Technological-Legal Lock-In' (2019) 119(7) Columbia Law Review 233

17. Darlington K, 'Aspects of Intelligent Systems Explanation' (2013) 1(2) Universal Journal of Control and Automation 40. doi:10.13189/ujca.2013.010204

18. Ferrante M, 'Causation in Criminal Responsibility' (2008) 11(3) New Criminal Law Review 470. doi:10.1525/nclr.2008.11.3.470

19. Griffiths O and Thorwart A, 'Effects of Outcome Predictability on Human Learning' (2017) 8 Frontiers in Psychology 511. doi:10.3389/fpsyg.2017.00511

20. Hallevy G, 'The Criminal Liability of Artificial Intelligence Entities-From Science Fiction to Legal Social Control' (2010) 4(2) Akron Intellectual Property Journal 171.

21. Hongladarom S and Van der Vaeren A, 'ChatGPT, Postphenomenology, and the Human-Technology-Nature Relations' (2024) 2 Journal of Human-Technology Relations 1. doi:10.59490/jhtr.2024.2.7386

22. Johnson GM, 'Algorithmic Bias: On the Implicit Biases of Social Technology' (2021) 198(10) Synthese 9941. doi:10.1007/s11229-020-02696-y

23. Kalogiannidis S and others, 'The Role of Artificial Intelligence Technology in Predictive Risk Assessment for Business Continuity: A Case Study of Greece' (2024) 12(2) Risks 19. doi:10.3390/risks12020019

24. Katyal SK, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66(1) UCLA Law Review 54.

25. Kuzma J and others, 'An Integrated Approach to Oversight Assessment for Emerging Technologies' in Marchant GE and Wallach W (eds), *Emerging Technologies: Ethics, Law and Governance* (Routledge 2020) 1197. doi:10.1111/j.1539-6924.2008.01086.x

26. Martins HMG, 'Liability Implications of Artificial Intelligence use in Health: Fault and Risk in Public Sector Healthcare' (Master's thesis, Universidade Catolica Portuguesa 2020)

27. Mineur C, 'Autonomous AI Technology and the Evolution of Legal Personhood in Criminal Law' (Master's thesis, University College Tilburg 2024)

28. Muller S, 'Visual Silence in the Language Portrait: Analyzing young People's Representations of their Linguistic Repertoires' (2022) 25(10) International Journal of Bilingual Education and Bilingualism 3644. doi:10.1080/13670050.2022.2072170

29. Navaneeth M, 'The Need for A Global Regulatory Framework for Artificial Intelligence: Implications of the European Union European Union Artificial Intelligence Act 2024' (Master's thesis, National University of Advanced Legal Studies 2024)

30. Neri R, 'Judging Beyond any Reasonable Doubt: A Logic and Epistemological Rule' (2024) 7 Quaestio Facti: Revista internacional sobre razonamiento probatorio 43. doi:10.33115/udg_bib/qf.i7.23028

31. Osmani N, 'The Complexity of Criminal Liability of AI Systems' (2020) 14(1) Masaryk University Journal of Law and Technology 53. doi:10.5817/mujlt2020-1-3

32. Oviedo-Trespalacios O and others, 'The Risks of Using ChatGPT to Obtain Common Safety-Related Information and Advice' (2023) 167 Safety Science 106244. doi:10.1016/j.ssci.2023.106244

33. Scherer M, 'Artificial Intelligence and Legal Decision-Making: The Wide Open?' (2019) 36(5) Journal of international arbitration 539. doi:10.54648/joia2019028

34. Simmler M, 'Responsibility Gap or Responsibility Shift? The Attribution of Criminal Responsibility in Human-Machine Interaction' (2024) 27(6) Information, Communication & Society 1142. doi:10.1080/1369118X.2023.2239895

35. Sison AJG and others, 'ChatGPT: More than a "Weapon of Mass Deception" Ethical Challenges and Responses from the Human-Centered Artificial Intelligence (HCAI) Perspective' (2024) 40(17) International Journal of Human-Computer Interaction 4853. doi:10.1080/10447318.2023.2225931

36. Solum LB, 'Legal Personhood for Artificial Intelligences' (1992) 70 North Carolina Law Review 1231

37. Stein AL, 'Assuming the Risks of Artificial Intelligence' (2022) 102 Boston University Law Review 979

38. Terzidou K, 'Generative AI for the Legal Profession: Facing the Implications of the Use of ChatGPT Through an Intradisciplinary Approach' (*MediaLaws*, 8 September 2023)

39. Turner J, 'Legal Personality for AI' in Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Springer 2018) 173. doi:10.1007/978-3-319-96235-1_5

40. Vaassen B, 'AI, Opacity, and Personal Autonomy' (2022) 35(4) Philosophy & Technology 88. doi:10.1007/s13347-022-00577-5

41. Wang J and others, 'Network Meets ChatGPT: Intent Autonomous Management, Control and Operation' (2023) 8(3) Journal of Communications and Information Networks 239. doi:10.23919/JCIN.2023.10272352

42. Wang X and Wu YC, 'Balancing Innovation and Regulation in the Age of Generative Artificial Intelligence' (2024) 14 Journal of Information Policy 385. doi:10.5325/jinfopoli.14.2024.0012

43. Wang Y and others, 'A Survey on ChatGPT: AI–Generated Contents, Challenges, and Solutions' (2023) 4 IEEE Open Journal of the Computer Society 280. doi:10.48550/arXiv.2305.18339

44. White F, 'Directive 85/374/EEC Concerning Liability for Defective Products: In the Name of Harmonisation, the Internal Market and Consumer Protection' in Giliker P (ed), *Research Handbook on EU Tort Law* (Edward Elgar 2017) 128. doi:10.4337/9781785365720.00013

45. Yan WJ and others, 'Navigating Uncertainties in Machine Learning for Structural Dynamics: A Comprehensive Review of Probabilistic and Non-Probabilistic Approaches in Forward and Inverse Problems' (*arXiv*:2408.08629, 16 August 2024). doi:10.48550/arXiv.2408.08629

46. Youvan DC, 'Reconciling the Probabilistic and Deterministic: Exploring Complexity, Emergence, and Uncertainty in Nature, AI, and Human Cognition' (*Research Gate*, October 2024). doi:10.13140/RG.2.2.16020.10880

47. Yuan Y and others, 'Does ChatGPT Know that it Does Not Know? Evaluating The Black-Box Calibration of Chatgpt' (2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), Torino, Italia, 20-25 May 2024) 5191.

48. Zetzsche DA, Buckley RP and Arner DW, 'The Distributed Liability of Distributed Ledgers: Legal Risks of Blockchain' (2018) 4 University of Illinois Law Review 1361. doi:10.2139/ssrn.3018214

AUTHORS INFORMATION

Raed S A Faqir
PhD, Associate Professor, Criminal Law, College of Law, American University in the Emirates, Dubai, United Arab Emirates.
raed.faqir@aue.ae
Associate Professor in Criminal Law, Faculty of Law, Al-Balqa Applied University, Al-Salt, Jordan,
r.faqir@bau.edu.jo
https://orcid.org/0000-0002-6102-0983
Corresponding author, solely responsible for the manuscript preparing.

Competing interests: No competing interests were disclosed.

Disclaimer: The author declares that his opinion and views expressed in this manuscript are free of any impact of any organizations.

RIGHTS AND PERMISSIONS

EDITORS

ABOUT THIS ARTICLE

Cite this article
Faqir RSA, 'ChatGPT in the Dock: Reflections on the Future of Criminal Liability' (2025) 8(Spec) Access to Justice in Eastern Europe 312-38 <https://doi.org/10.33327/AJEE-18-8.S-o000156>

DOI: https://doi.org/10.33327/AJEE-18-8.S-o000156

Keywords: *generative artificial intelligence, ChatGPT, criminal liability, artificial actor, AI act, legal innovation, algorithmic risk.*

DETAILS FOR PUBLICATION

Date of submission: 19 Aug 2025
Date of acceptance: 30 Oct 2025
Date of Publication: 30 Dec 2025
Whether the manuscript was fast tracked? - Yes
Number of reviewer report submitted in first round: 2 reports
Number of revision rounds: 1 round with conditionally acceptance

Technical tools were used in the editorial process:
Plagiarism checks - Turnitin from iThenticate https://www.turnitin.com/products/ithenticate/
Scholastica for Peer Review https://scholasticahq.com/law-reviews

AI DISCLOSURE STATEMENT

I confirm that no artificial intelligence tools or services were used at any stage of writing, translating, editing, or analyzing content for this manuscript.

Access to Justice in Eastern Europe
ISSN 2663-0575 (Print) ISSN 2663-0583 (Online)
Journal homepage http://ajee-journal.com

АНОТАЦІЯ УКРАЇНСЬКОЮ МОВОЮ

Стаття-думка

CHATGPT НА ЛАВІ ПІДСУДНИХ: РОЗДУМИ ПРО МАЙБУТНЄ ІНСТИТУТУ КРИМІНАЛЬНОЇ ВІДПОВІДАЛЬНОСТІ

Раед С. А. Факір

АНОТАЦІЯ

Вступ. *У цьому дослідженні розглядаються правові проблеми, що виникають через генеративний штучний інтелект. Також було підкреслено обмеження традиційної системи інституту кримінальної відповідальності щодо шкоди, заподіяної результатами роботи ШІ. Дослідження вивчає нові моделі відповідальності для забезпечення підзвітності та захисту прав особи в епоху інтелектуальних машин.*

Генеративний штучний інтелект (ШІ), прикладом якого є ChatGPT, перетворився з простого обчислювального інструменту на когнітивного агента, здатного створювати контент, вирішувати проблеми та ухвалювати рішення. Ця еволюція кидає виклик традиційній системі інституту кримінального права, піднімаючи складні питання про притягнення до відповідальності, у випадках, коли результати, згенеровані ШІ, завдають шкоди або призводять до злочинної поведінки. Дослідження «ChatGPT на лаві підсудних: роздуми про майбутнє інституту кримінальної відповідальності» розглядає ці правові дилеми, спричинені втручанням штучного інтелекту, зосереджуючись на недоліках традиційних концепцій кримінальної відповідальності та вивчаючи потребу в нових правових парадигмах.

Методи. *У дослідженні використано описово-аналітичну та порівняльну методику. Автор здійснив аналіз національного та міжнародного законодавства, правових принципів та сучасної юриспруденції, зосередивши увагу на Європейському законі про штучний інтелект (2024) як моделі. Дослідження вивчає автономні можливості штучного інтелекту, непрозорість алгоритмічного ухвалення рішень і труднощі встановлення причинно-наслідкових зв'язків між діями штучного інтелекту та завданою шкодою. Тематичні дослідження використовуються для вивчення потенційних моделей відповідальності, зокрема превентивної відповідальності та концепції «штучного актора».*

Результати та висновки. *У статті було з'ясовано, що традиційна система інституту кримінальної відповідальності є невідповідною для таких систем штучного інтелекту, як ChatGPT, з огляду на їх часткову автономію та алгоритмічну складність. Також було виявлено потенціал для розширення відповідальності розробників, операторів і користувачів, а також необхідність гнучких правових моделей, які поєднують превентивні, адміністративні та кримінальні заходи. Дослідження підкреслює важливість інтеграції правових інновацій із технологічним наглядом для захисту прав людини, зберігаючи при цьому запобіжну та захисну функції кримінального права.*

Ключові слова: *генеративний штучний інтелект, ChatGPT, кримінальна відповідальність, штучний актор, закон про ШІ, правова інновація, алгоритмічний ризик.*