



QUALIFICATION OF CRIMES COMMITTED USING ARTIFICIAL INTELLIGENCE: ISSUES OF CRIMINAL LAW ASSESSMENT

Albina Kurmichkina

Lecturer of Cyber Law Department, Tashkent State University of Law,
Uzbekistan

albinaakurmichkina@gmail.com

0009-0000-6758-6227

<https://doi.org/10.5281/zenodo.17809216>

ARTICLE INFO

Received: 27th November 2025

Accepted: 28th November 2025

Online: 29th November 2025

KEYWORDS

The problematic nature of qualifying crimes involving AI extends beyond mere technological novelty

ABSTRACT

The exponential advancement of artificial intelligence technologies has precipitated unprecedented challenges to established criminal law doctrines, fundamentally questioning the conceptual foundations upon which criminal liability has been traditionally constructed. The integration of autonomous and semi-autonomous AI systems into various spheres of human activity necessitates a comprehensive reexamination of conventional approaches to criminal qualification, particularly regarding the attribution of mens rea, the establishment of causal connections, and the identification of appropriate subjects of criminal responsibility [1]. Contemporary legal scholarship has increasingly recognized that existing criminal law frameworks, predicated upon anthropocentric conceptions of agency and intentionality, prove inadequate when confronted with the operational realities of machine learning algorithms and autonomous decision-making systems [2].

Introduction

The exponential advancement of artificial intelligence technologies has precipitated unprecedented challenges to established criminal law doctrines, fundamentally questioning the conceptual foundations upon which criminal liability has been traditionally constructed. The integration of autonomous and semi-autonomous AI systems into various spheres of human activity necessitates a comprehensive reexamination of conventional approaches to criminal qualification, particularly regarding the attribution of *mens rea*, the establishment of causal connections, and the identification of appropriate subjects of criminal responsibility [1]. Contemporary legal scholarship has increasingly recognized that existing criminal law frameworks, predicated upon anthropocentric conceptions of agency and intentionality, prove inadequate when confronted with the operational realities of machine learning algorithms and autonomous decision-making systems [2].

The problematic nature of qualifying crimes involving AI extends beyond mere technological novelty; it strikes at the epistemological core of criminal jurisprudence itself. Traditional criminal law theory presupposes a rational human actor whose conscious volition

constitutes the essential element of culpability [3]. However, the emergence of AI systems capable of independent learning, pattern recognition, and decision-making without direct human intervention has rendered this foundational assumption increasingly tenuous. The doctrinal difficulty becomes particularly acute when considering scenarios wherein AI systems generate outcomes that, had they been produced by human actors, would unquestionably constitute criminal conduct, yet the relationship between human programming, algorithmic autonomy, and harmful consequences remains opaque and multifaceted [4].

The significance of this inquiry transcends theoretical discourse, bearing profound implications for the practical administration of criminal justice in an increasingly digitalized society. As jurisdictions worldwide grapple with crimes facilitated or perpetrated through AI mechanisms—ranging from algorithmic fraud and automated market manipulation to AI-enabled surveillance violations and autonomous vehicle fatalities—the urgent necessity for coherent qualification frameworks becomes manifest [5]. The present analysis endeavors to illuminate the principal doctrinal tensions inherent in applying traditional criminal law concepts to AI-mediated conduct, while simultaneously exploring potential theoretical reconstructions that might accommodate the unique characteristics of algorithmic agency.

Methodology

This doctrinal research employs a multidimensional analytical framework, synthesizing comparative legal analysis with theoretical criminological inquiry to examine the qualification challenges posed by AI-involved criminality. The methodological approach encompasses three interconnected analytical dimensions: conceptual analysis of fundamental criminal law principles, comparative examination of emerging legislative and judicial responses across multiple jurisdictions, and theoretical exploration of potential doctrinal adaptations.

The conceptual analysis interrogates the foundational elements of criminal liability—*actus reus*, *mens rea*, and causation—through the prism of AI functionality, examining the extent to which these traditional constructs retain analytical utility when applied to algorithmic conduct. This inquiry necessarily engages with both Anglo-American and Continental European legal traditions, given their divergent approaches to criminal culpability and their varying degrees of flexibility in accommodating technological innovation [6]. The comparative dimension surveys legislative initiatives and judicial decisions from jurisdictions at the forefront of AI regulation, including the European Union, United States, United Kingdom, and selected Asian jurisdictions, to identify emergent patterns in legal responses to AI-mediated harm [7].

The theoretical component draws upon interdisciplinary insights from computer science, philosophy of law, and criminology to construct a comprehensive understanding of how algorithmic operations intersect with criminal law doctrine. Particular attention is devoted to the taxonomy of AI involvement in criminal conduct, distinguishing between AI as instrumentality, AI as autonomous agent, and hybrid scenarios wherein human and algorithmic decision-making prove inseparably intertwined. This taxonomical approach facilitates more nuanced analysis of qualification issues, recognizing that different modalities of AI involvement present distinct doctrinal challenges requiring tailored legal responses [8].

Results

The analysis reveals several critical fault lines in the application of traditional criminal qualification frameworks to AI-mediated conduct, each presenting distinct theoretical and

practical challenges to legal practitioners and policymakers. The first and perhaps most fundamental difficulty concerns the attribution of criminal intent in circumstances where harmful outcomes result from algorithmic processes whose operational logic may be opaque even to their creators. The classical requirement of *mens rea* presupposes a mental state capable of forming criminal intent, yet AI systems, regardless of their sophistication, lack consciousness in any meaningful sense [9]. This creates what may be termed the "intent attribution gap"—a conceptual space wherein harmful conduct occurs with neither clear human intentionality nor appropriate legal mechanisms for attribution.

The doctrine of strict liability, while offering a potential pathway for circumventing *mens rea* requirements in certain regulatory contexts, proves insufficient as a comprehensive solution. Strict liability offenses traditionally operate within narrowly circumscribed domains and typically involve relatively minor penalties, reflecting the reduced culpability associated with absence of intent [10]. Extending this approach to serious AI-mediated harms would represent a radical departure from fundamental principles of criminal justice, potentially imposing severe sanctions absent the moral opprobrium that traditionally justifies criminal punishment. Moreover, the indiscriminate application of strict liability to AI-related conduct risks creating perverse incentives, potentially discouraging beneficial innovation while failing to adequately address genuinely culpable behavior [11].

The second critical finding concerns the problematic nature of establishing causal links between human conduct and AI-generated harm. Traditional causation doctrine employs frameworks such as "but-for" causation and proximate cause to delineate the boundaries of criminal responsibility, predicated upon relatively straightforward causal chains linking human action to prohibited consequences [12]. However, AI systems, particularly those employing machine learning methodologies, operate through iterative processes of pattern recognition and prediction that may generate outcomes substantially divergent from programmer intentions or reasonably foreseeable consequences. The "black box" problem—wherein even sophisticated technical analysis cannot fully explicate how particular algorithmic decisions emerge from training data and neural network operations—fundamentally complicates traditional causal analysis [13].

This causal indeterminacy becomes especially pronounced in scenarios involving autonomous learning systems that modify their decision-making parameters through interaction with operational environments. When an AI system, initially programmed for legitimate purposes, evolves through unsupervised learning to generate harmful outcomes, the attribution of responsibility becomes extraordinarily complex. The temporal and conceptual distance between initial programming and ultimate harm may be so attenuated that conventional causation frameworks prove analytically inadequate [14]. Some scholars have proposed supplementing traditional causation analysis with concepts borrowed from product liability law, treating AI systems as potentially defective products subject to special liability regimes, though this approach raises its own doctrinal difficulties regarding the appropriateness of civil law concepts in criminal contexts [15].

The third significant finding relates to the fundamental question of identifying appropriate subjects of criminal responsibility in AI-mediated offenses. The principle of personal culpability, central to modern criminal law, presumes identifiable human actors whose conduct and mental

state satisfy the elements of offense definitions [16]. However, the distributed nature of AI development and deployment—involving programmers, training data curators, deployment entities, and end users—fragments responsibility across multiple actors whose individual contributions may be individually insufficient to establish criminal liability under traditional doctrines of perpetration or complicity. The emergence of "collective algorithmic responsibility" wherein harm results from the cumulative effect of numerous individually innocuous decisions presents profound challenges to individualistic criminal law frameworks [17].

Existing doctrines of corporate criminal liability offer limited guidance for addressing this issue. While corporate criminal responsibility recognizes that organizational entities may be held accountable for crimes committed by agents acting within the scope of employment, these frameworks remain fundamentally predicated upon attributing human conduct to juridical persons [18]. The question of whether AI systems themselves might constitute appropriate subjects of criminal liability—a proposition that would require radical reconceptualization of fundamental criminal law principles—remains deeply controversial within legal scholarship, with most commentators rejecting such an approach as conceptually incoherent given AI's lack of moral agency [19].

Discussion

The preceding analysis demonstrates that qualification of AI-mediated crimes cannot be adequately addressed through minor adjustments to existing criminal law frameworks; rather, it necessitates fundamental theoretical reconstruction of core criminal law concepts. The path forward requires careful navigation between competing imperatives: maintaining criminal law's foundational commitment to personal culpability and moral desert while simultaneously ensuring that legal frameworks remain capable of addressing genuinely harmful conduct facilitated by technological advancement.

One promising theoretical direction involves developing graduated liability frameworks that calibrate criminal responsibility according to the degree of human control, foreseeability, and intervention capability present in specific AI deployment contexts. Such frameworks would distinguish between scenarios involving fully autonomous AI systems operating beyond meaningful human control and situations wherein AI functions as a sophisticated tool subject to human direction and oversight [20]. This graduated approach recognizes that AI involvement exists along a spectrum rather than constituting a binary category, permitting more nuanced legal responses attuned to varying degrees of human agency and culpability.

The doctrine of "algorithmic negligence" represents another potentially fruitful avenue for doctrinal development. This concept would impose duties of care upon those who develop, deploy, or utilize AI systems, with criminal liability attaching to failures to meet requisite standards of precaution [21]. Such an approach builds upon existing negligence-based criminal offenses while adapting them to accommodate the unique risks posed by algorithmic decision-making. Critical to this framework's success would be the articulation of clear standards regarding reasonable algorithmic design, testing, monitoring, and intervention protocols, potentially drawing upon emerging technical standards and best practices within the computer science community.

The comparative analysis reveals instructive divergence in jurisdictional approaches to these challenges. The European Union's AI Act, while primarily regulatory rather than criminal in

orientation, establishes a risk-based classification system that could inform criminal law qualification by distinguishing between AI applications presenting varying levels of potential harm [22]. Conversely, certain U.S. jurisdictions have begun exploring expanded application of existing criminal statutes to AI contexts, testing the flexibility of traditional doctrines without comprehensive legislative reform. Each approach presents distinct advantages and limitations, suggesting that optimal legal frameworks may incorporate elements of both regulatory classification and common law adaptability.

The question of establishing specialized criminal offenses specifically addressing AI-mediated harm merits serious consideration. Such offenses might criminalize particular conduct antecedent to harm production—such as deployment of inadequately tested AI systems in high-risk contexts or failure to implement required safety mechanisms—thereby circumventing difficulties in establishing traditional *mens rea* while maintaining meaningful connection between criminal liability and culpable conduct [23]. However, the proliferation of specialized AI-related offenses risks creating fragmented and incoherent criminal codes, potentially undermining principles of legal certainty and fair notice essential to legitimate criminal justice systems.

International cooperation emerges as essential to effective legal responses to AI-mediated crime, given both the transnational nature of AI development and the ease with which digital technologies transcend territorial boundaries. The absence of harmonized approaches to criminal qualification creates opportunities for regulatory arbitrage and complicates cross-border law enforcement cooperation [24]. International instruments establishing minimum standards for AI-related criminal liability, while respecting legitimate differences in national legal traditions, could significantly enhance the effectiveness of domestic criminal justice responses while promoting consistency and predictability in this rapidly evolving legal domain.

Conclusion

The qualification of crimes involving artificial intelligence represents one of contemporary criminal law's most formidable challenges, requiring nothing less than fundamental reconceptualization of traditional doctrines to accommodate the realities of algorithmic agency. The analysis presented herein demonstrates that neither wholesale abandonment of established criminal law principles nor their mechanical application to novel technological contexts offers satisfactory resolution. Rather, the path forward demands creative theoretical synthesis, developing frameworks that preserve criminal law's essential normative commitments while adapting its technical machinery to address genuinely harmful conduct in an increasingly AI-mediated world. As AI technologies continue their inexorable advance, the urgency of this doctrinal reconstruction intensifies, demanding sustained engagement from legal scholars, practitioners, and policymakers committed to ensuring that criminal law remains an effective instrument of justice in the twenty-first century.

References:

- [1] Hallevy, G. (2015). *Liability for Crimes Involving Artificial Intelligence Systems*. Springer.
- [2] Hildebrandt, M. (2018). Law as computation in the era of artificial legal intelligence: Speaking law to the power of statistics. *University of Toronto Law Journal*, 68(1), 12-35.

- [3] Hart, H.L.A. (2008). *Punishment and Responsibility: Essays in the Philosophy of Law* (2nd ed.). Oxford University Press.
- [4] Asaro, P. (2016). The liability problem for autonomous artificial agents. *Proceedings of the AAAI Spring Symposium on Ethical and Moral Considerations in Non-Human Agents*.
- [5] Yeung, K., & Lodge, M. (2019). *Algorithmic Regulation*. Oxford University Press.
- [6] Dubber, M.D. (2005). Comparative Criminal Law. In M. Reimann & R. Zimmermann (Eds.), *The Oxford Handbook of Comparative Law* (pp. 1287-1309). Oxford University Press.
- [7] Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.
- [8] Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), 175-183.
- [9] Searle, J.R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424.
- [10] Simester, A.P., & Sullivan, G.R. (2010). *Criminal Law: Theory and Doctrine* (4th ed.). Hart Publishing.
- [11] Lample, G., & Charton, F. (2020). Deep Learning for Symbolic Mathematics. *International Conference on Learning Representations*.
- [12] Moore, M.S. (2009). *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*. Oxford University Press.
- [13] Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1-12.
- [14] Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21.
- [15] Schellekens, M. (2015). Self-driving cars and the chilling effect of liability law. *Computer Law & Security Review*, 31(4), 506-517.
- [16] Ashworth, A., & Horder, J. (2013). *Principles of Criminal Law* (7th ed.). Oxford University Press.
- [17] Danaher, J. (2016). Robots, law and the retribution gap. *Ethics and Information Technology*, 18(4), 299-309.
- [18] Wells, C. (2001). *Corporations and Criminal Responsibility* (2nd ed.). Oxford University Press.
- [19] Pagallo, U. (2013). *The Laws of Robots: Crimes, Contracts, and Torts*. Springer.
- [20] Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review*, 103(3), 513-563.
- [21] Abbott, R. (2020). The Reasonable Computer: Disrupting the Paradigm of Tort Liability. *George Washington Law Review*, 86(1), 1-45.
- [22] European Commission. (2021). *The EU Artificial Intelligence Act*.
- [23] Kingston, J. (2016). Artificial Intelligence and Legal Liability. In M. Bramer & M. Petridis (Eds.), *Research and Development in Intelligent Systems XXXIII* (pp. 269-279). Springer.
- [24] Ryngaert, C., & Taylor, M. (2020). The GDPR as global data protection regulation? *American Journal of International Law Unbound*, 114, 5-9.