Contemporary Challenges: The Global Crime and Security Journal (2025) Vol. 6

Research Article

Collective Risk Allocation and Restorative Justice in the Age of Artificial Intelligence

Dr. Maria Carla Canato

Received: 28 January 2025 / Accepted: 10 June 2025

Abstract

This paper aims to reflect on the issue of the 'responsibility gap' introduced by artificial intelligence (AI) within the criminal law context. Starting from this critical point, it outlines potential pathways to move beyond traditional punitive models, placing particular emphasis on the relevance of restorative approaches. The first part highlights the challenges that AI poses to the applicability of conventional criminal law institutions. The analysis then shifts to the mechanisms through which risk associated with the use of AI is distributed among the various actors involved in its design and deployment, underlining not only the importance of accountability, but above all the systemic and collective nature of 'AI-related risk'.

Within this framework, the concluding section explores restorative justice as a possible tool to address the consequences of algorithmic harm, emphasising mechanisms grounded in dialogue, trust-building, and social responsibility. The core thesis advanced is the need to rethink the notions of culpability and responsibility in the AI era, moving beyond the primacy of the individual toward a more systemic and shared perspective. In this sense, the paper aims to make a critical contribution to the development of a justice model that addresses new digital challenges through the lens of distributed and collective responsibility, while ensuring that the protection of fundamental rights, particularly in the criminal domain, continues to affirm its inherently human dimension.

Introduction

The use of artificial intelligence (AI) has impacted various sectors, including healthcare, transportation, and criminal justice, raising significant ethical, legal, and philosophical discussions. The primary issue associated with the use of AI relates to what has been termed in the literature as the 'responsibility gap', understood as the situation in which it is challenging to identify the party accountable for the consequences arising from the deployment of AI systems.[289] Given the influence of such tools in sensitive decision-making contexts — such as those connected to healthcare[290], the judiciary[291], or autonomous vehicle operation[292] — the limitations of applying traditional liability models become clear[293]. The opacity of AI algorithms, their capacity for self-learning, and their reliance on vast datasets[294] further complicate the attribution of responsibility.[295]

The rapid pace of technological innovation, especially in artificial intelligence, has outstripped the evolution of legal and ethical frameworks, resulting in a significant 'responsibility gap'[296] that challenges the core principles of criminal law. This article aims to examine the impact of the so-called responsibility gap on traditional liability frameworks, while exploring alternative models of justice — specifically restorative justice — as a potential response to the complexities of the digital age. The proposed model shifts the focus away from individual liability to investigate the role of collective responsibility[297], which redistributes the risk associated with artificial intelligence across all stages of AI system development. It conceptualises responsibility within a

[289] Matthias, A., *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, in *Ethics and Information Technology,* vol. 6, (2004): 1 et seq.

[290] Bartlett, A., *The possibility of AI-induced medical manslaughter: Unexplainable decisions, epistemic vices, and a new dimension of moral luck*, in *Medical Law International* (2023).

[291] Darshan, V. *Demystifying the Role of Artificial Intelligence in Legal Practice*, in *Nirma University Law Journal*, vol. 8, no.2 (2019).

[292] Hilgendord, E., *Automated Driving and the Law*, in *Robotics, Autonomics and the Law*, (E. Hilgendorf-U. Seidel, 2017): 181-182.

[293] On the limits of criminal liability, see recently Fragasso, B., *Intelligenza artificiale e responsabilità penale: principi e categorie alla prova di una tecnologia "imprevedibile"*. (Turin: Giappichelli, 2025).

[294] Doncieux, S., Mouret, J.B., *Beyond black-box optimization: a review of selective pressures for evolutionary robotics*, in *Evolutionary Intelligence* (2014): 71 et seq.

[295] Salvadori, I., *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale, RIDPP* (2021): 90 et seq.

[296] Matthias, A., *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata, op. cit.*

[297] Taylor, I. *Collective Responsibility and Artificial Intelligence*, in *Philos. Technol.* Vol. 37, n. 27 (2024).

multi-layered framework that encompasses all actors involved in the design, implementation, and use of such technologies.[298]

Given the well-documented complexity and opacity of the aforementioned systems, as well as the large number of actors involved in their use[299], it often becomes unfeasible to allocate liability for harms to individual persons. This implies the need to establish a new model of shared responsibility. In this respect — and in light of the inherent irreducibility of AI-related risks, as acknowledged by the risk-based approach adopted in the AI Act[300] — it appears promising to respond through the institutionalisation of non-criminal strict liability regimes for violations of technical, ethical, and safety standards, within a broader framework of accountability.[301]

Such mechanisms enable the alignment of AI systems with relevant standards, promoting effective risk management without resorting to criminal proceedings. Criminal law, in fact, proves inadequate not only in terms of timeliness but also, and often more critically, in ensuring the effective protection of fundamental rights. This is due to the inherent difficulty — if not impossibility — of proving guilt beyond a reasonable doubt in many cases.

Therefore, in the wake of enhanced accountability frameworks, the valorisation of non-criminal restorative approaches — such as restorative justice — offers a human-centred method for addressing the negative outcomes stemming from the use of artificial intelligence. This model emphasises dialogue, repair, and social reintegration over punitive responses, thereby seeking to fulfil the rehabilitative function of criminal law.

Ultimately, this integrated governance framework places the effective protection of

---

[298] Lima, D., *Could AI Agents Be Held Criminally Liable: Artificial Intelligence and the Challenges for Criminal Law*, in *South Carolina Law Review*, vol. 69 (2018): 687. In addressing the possible analogies between AI and legal entities, it is noted that legal entities do not exist autonomously, as they are closely tied to individuals and represent an artificial construct created by them. In contrast, AI systems seem to surpass this artificial nature, as they are increasingly capable of operating without the need for direct human intervention.

[299] Consulich, F., *Il concorso di persone nel reato colposo* (Torino: 2023), 1 ss.

[300] European Union. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) 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*. Official Journal of the European Union, L 259 (July 12, 2024): 1–294.

[301] Panattoni, B., *AI and Criminal Law: The Myth of 'Control' in a Data-Driven Society*, in *Revue Internationale de Droit Pénal*, (2021): 138.

fundamental rights at the core of the balance between technological risk, ethical imperatives, and the foundational principles of criminal law.

Within this context, the concept of 'permitted risk'[302] emerges as a key legal instrument to address and resolve issues of liability in AI use. Traditionally, this concept has been employed in criminal law to delineate the boundaries of culpability in situations where certain residual risks persist despite the adoption of all mandated precautions.

The idea of permitted risk acknowledges that some level of danger is inherent to complex or innovative processes, and the legal system has typically drawn a distinction between acceptable risks and those that cross the threshold of criminal responsibility. However, with the rapid advancements in AI, these boundaries become increasingly challenging to define, as the systems driving AI technologies often operate in ways that are not fully predictable, even by experts.[303]

AI systems introduce new forms of risk, especially risks that arise from their ability to learn, adapt, and make decisions independently of direct human control.[304] These characteristics challenge traditional legal frameworks, where responsibility is typically assigned based on human intention or negligence.[305]

The nature of AI, however, complicates this by blurring the lines between human agency and machine autonomy. When an AI system causes harm or damage, it is often unclear who, if anyone, should be held accountable: the developers who programmed the AI, the users who deployed it, or the AI itself, as an independent actor? In this sense, the concept of permitted risk requires adaptation.[306] While the

---

[302] Consulich, F., *Rischio consentito*, in *Reato colposo*, *Enciclopedia del Diritto*, ed. M. Donini (Milano: Giuffrè, 2021).

[303] Gless, S., Silverman, E., Weigend, T., *If robots cause harm, who is to blame? Self-driving cars and criminal liability*, *New Criminal Law Review*, (2016): 425 et seq.

[304] On this point, see Doncieux, S., Mouret, J.B., *Beyond black-box optimization: a review of selective pressures for evolutionary robotics, op. cit*. (2014): 71 et seq.

[305] Pagallo, U., *The Laws of Robots. Crimes, Contracts and Torts* (Springer, 2013), 47; Beck, S., *Google Cars, Software Agents, Autonomous Weapons Systems – New Challenges for Criminal Law?*, *Law, Computer Science* (2017): 243. Regarding autonomous cars, see Surden, H., Williams, M.A., *Technological Opacity, Predictability, and Self-Driving Cars*, in *Cardozo Law Review* (2016): 157 et seq.

[306] Although unrelated to artificial intelligence, recent European jurisprudence underscores the structural limitations of assigning legal responsibility solely to individual agents in cases of diffuse, cumulative, or probabilistic harm. The shift evident in these rulings points toward a normative logic of *precautionary accountability*—a forward-looking responsibility model that prioritizes institutional vigilance, shared risk management, and systemic safeguards over retrospective fault attribution. This orientation, grounded in environmental and human rights law, offers a valuable precedent for AI governance, where causal opacity and collective agency are likewise central. See:

legal system may acknowledge that certain AI-driven risks are foreseeable, it must also consider whether the systems in question were developed or used in a manner that reasonably minimised those risks. Furthermore, it raises questions about the ethical responsibility of those who deploy AI technologies without fully understanding or mitigating their potential consequences.

This perspective naturally leads to the exploration of 'collective responsibility' as a way to address the systemic nature of AI risks. AI development and deployment are not solely individual efforts; they involve a wide array of stakeholders, including developers, corporations, regulators, and users.[307] As such, the ethical and legal implications of AI technologies cannot be contained within the scope of individual actions. Instead, responsibility must be distributed across the entire ecosystem that enables and sustains AI[308].

I. Precautionary Accountability, the AI Act, and Restorative Justice

The concept of collective responsibility emphasises the need for coordinated and multi-stakeholder efforts to ensure that AI systems are developed, deployed, and operated with due diligence, transparency, and adherence to robust ethical standards. While the AI Act does not explicitly invoke the terminology of 'collective responsibility', its regulatory framework implicitly fosters this approach by imposing comprehensive obligations on a wide range of actors, particularly enterprises, to uphold stringent requirements throughout the AI lifecycle[309].

– *Urgenda Foundation v. State of the Netherlands*, Supreme Court of the Netherlands, Judgment of 20 December 2019, ECLI:NL:HR:2019:2007 (affirming the state's tort liability for climate inaction based on positive duties of prevention).

– *KlimaSeniorinnen v. Switzerland*, ECtHR, Grand Chamber, Judgment of 9 April 2024, App. No. 53600/20 (establishing a state obligation under Article 8 ECHR to prevent foreseeable harm caused by systemic environmental risk).

– *Lliuya v. RWE AG*, Oberlandesgericht Hamm (Germany), Case No. 5 U 15/17, appeal pending as of 2024 (exploring corporate tort liability for global emissions despite partial and indirect contribution).

– *Cannavacciuolo and Others v. Italy*, ECtHR, Chamber Judgment of January 2025, App. No. 43247/19 (expanding Article 2 ECHR to include deaths linked to long-term state inaction on environmental hazards).

[307] Fragasso, B., *La responsabilità penale del produttore di sistemi di intelligenza artificiale*, in *Sistema Penale*, (2023): 1 et seq.

[308] Van de Poel, I., Sand, M., *Varieties of responsibility: two problems of responsible innovation*, in *Synthese* 198 (Suppl 19), 4769–4787 (2021).

[309] As already noticed, recent evolutions in European jurisprudence suggest a growing recognition of the structural inadequacy of individual fault-based liability when dealing with distributed and emergent

By setting harmonised standards for risk management data governance, transparency, and human oversight, the AI Act creates a legal environment where accountability is distributed across various organisational levels and functions. This diffusion of responsibility encourages companies to embed ethical and safety considerations into their processes and governance structures, thereby contributing to a systemic model of accountability.[310] In this sense, the Act aligns with the broader notion that ensuring trustworthy AI is not solely the burden of individual developers or users but a shared enterprise involving designers, manufacturers, deployers, and regulators.

Furthermore, the AI Act's framework reflects an awareness of the 'permitted risk'[311], inherent to technological innovation. It implicitly acknowledges that while some degree of risk cannot be entirely eliminated, it must be proactively managed through preventive measures and compliance with established standards. This approach aligns with the paradigm of collective responsibility, emphasising that the duty to prevent harm does not rest with isolated individual agents, but rather extends to all actors within the digital ecosystem—including the political, corporate, and social domains.[312]

---

harms—particularly in contexts characterized by cumulative risk, epistemic uncertainty, and organizational complexity. This has prompted a normative turn toward precautionary accountability, a forward-looking model that emphasizes shared responsibility, institutional oversight, and ex ante harm mitigation over traditional ex post blame assignment. Such a shift aligns with several core provisions of the AI Act, which introduces a risk-based regulatory framework that imposes preventive duties—e.g., risk assessments, traceability, human oversight—not on individuals per se, but on systemic actors such as providers, deployers, and notified bodies. These duties are expressly designed to operate irrespective of individual foreseeability or intent, reflecting the epistemic and causal fragmentation typical of AI systems.mechanisms, and forward-oriented accountability processes, complementing the AI Act's preventive ethos and bridging the residual responsibility gaps.

[310] Giannini, A., '*Criminal Behavior and Accountability of Artificial Intelligence Systems'* (Maastricht Law Series, 2023), 180–185.

[311] In fact, the notion of permitted risk has two possible interpretations: one that associates it with all dangerous but socially accepted activities, and the other that uses it to define the residual risk that remains after all possible precautions have been taken and is therefore accepted simply because it is unavoidable, provided that the legal norm does not prohibit it following a cost-benefit balance (here, the purpose of the violated norm may have potential utility). On the dual meanings, see Forti, G., *Colpa ed evento nel diritto penale*, (Giuffré, 1990), 456. On the concept of permitted risk as the general risk of daily life or tolerated residual risk, see Frisch, W., *Zum gegenwärtigen Stand der Diskussion und zur Problematik der objektiven Zurechnungslehre*, *G.A. StR.* (2003): 723; on this matter, see also, Roxin, C., Greco, L., *Strafrecht, Allgemeiner Teil,* (C.H. Beck, 2020), 488.

[312] This anticipatory structure resonates with the jurisprudential reasoning in *Urgenda* and *KlimaSeniorinnen*, where courts imposed legal duties on states to act preventively in the face of probabilistic, yet foreseeable, collective harm. Similarly, the precautionary logic embedded in the AI Act

The relationship between the protection of fundamental rights and the prevention of systemic risk arising from AI, which underpins recent European legislation, aims to establish a governance model that is both adaptive and precautionary. It is essential to highlight that this system delineates a threshold between 'acceptable' and 'irreducible' risks, on the one hand, and 'unacceptable' risks, on the other.

Ultimately, the AI Act represents a crucial step toward the implementation, at the European level, of a form of collective accountability for artificial intelligence. This model is grounded in compliance with development and deployment standards and sustained by ongoing regulatory oversight. It complements — without displacing — traditional notions of individual liability.

As AI increasingly permeates sensitive contexts, the legal framework must evolve to balance technological innovation with ethical imperatives, ensuring that responsibility is appropriately allocated, based on clearly defined and anticipatory obligations, and oriented toward the protection of fundamental rights and social welfare.

This integrated, multilevel governance approach—including at the supranational level—highlights the need to reconceptualise regulation through the lens of equity.[313] It calls for a cost-benefit balance that, while recognising the shared nature of risks[314], restores a personalised perspective in the face of harm[315].

Within this complex framework,[316] restorative justice should be fully incorporated into AI governance, contributing to the realisation of 'collective' accountability through a dialogic approach centred on the individual, non-retributive, but rehabilitative and remedial in nature.

Ultimately, this discussion highlights the importance of a holistic and adaptive approach to justice, one that extends beyond traditional notions of individual

---

can be seen as prefiguring a framework of distributed, institutional responsibility: a form of collective duty of care suited to socio-technical environments.

[313] On this topic, see Novelli, C., *L'Artificial Intelligence Act Europeo: alcune questioni di implementazione*, *federalismi.it*, no.2 (2024): 95.

[314] Ortalda, A., De Hert, P., *Artificial Human Rights Impact Assessment*, in *Artificial Intelligence and Human Rights*, (A. Quintavalla and J. Temperman, 2023), 532. The authors argue that 'Not only technology changes. Human rights are equally undergoing a period of change'.

[315] Novelli, C., Casolari, F., Rotolo, A., Taddeo, M., Floridi, L., *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act*, in *Digital Society* (2024): 18 et seq.

[316] In fact, 'permitted risk' is not determined by looking at the type of danger itself but depends on the qualification of the activity from which it arises, in relation to the precautions that the legal system has imposed on it in light of considerations of general utility.

accountability to encompass collective responsibility and restorative justice. As AI continues to reshape our world, justice systems must evolve to ensure that the benefits of this transformative technology are balanced with the imperatives of accountability, fairness, and equity. The law must be equipped to address not only the risks associated with AI but also the profound ethical questions that arise when we relinquish some degree of control to intelligent systems.[317] In doing so, we can build a future where technological advancements do not outpace our moral and legal obligations, but rather operate within a framework of responsibility and care that reflects our shared values.

A. The Philosophy of Punishment in the Context of AI

Traditional theories of punishment, including retribution, deterrence, and rehabilitation, presuppose a moral agent capable of intentional action and understanding the consequences of their behaviour.[318] However, AI systems, particularly those based on machine learning, lack intentionality, moral awareness, and the capacity for guilt or remorse.[319] This absence of moral agency[320] undermines the applicability of these

[317] Freitas, P., Andrade, F., Novais, P., *Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible*, in Casanovas-Pagallo-Palmirani-Sartor, *AI Approaches to the Complexity of Legal Systems*, *AICOL-IV and AICOL-V International Workshops 2013*, (Springer, 2014), 145 et seq.

[318] Hallevy, G., *The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control*, in *Akron Intellectual Property Journal (*2010):171 et seq.

[319] Hallevy, G., *Liability for Crimes Involving Artificial Intelligence Systems*, (Springer, 2015), 47 et seq. According to the author, there are no valid reasons to deny the criminal liability of Artificial Intelligence (AI) systems.
Hallevy theorizes that advanced AI systems may fulfill both the *actus reus* and *mens rea* requirements of criminal law, given their capacity for autonomous conduct, data acquisition, and probabilistic reasoning. He contends that certain mental states—such as negligence, recklessness, or general intent—may be ascribed to such systems, despite their lack of consciousness or emotions. Rejecting the traditional axiom that machines cannot commit or be punished for crimes, Hallevy proposes granting AI legal personhood and articulates three models of liability: (1) perpetration via another, where the AI acts as an innocent agent under human control; (2) natural probable consequence, involving derivative or shared liability; and (3) direct liability, whereby the AI may be held independently accountable for its actions.

[320] Electronic personhood, as envisioned by its proponents, aims to balance the injured party's right to compensation—with access to indemnification that does not depend on proving a product defect or user fault—with the need to foster innovation among businesses and to ensure predictability in judicial decisions. In particular, assigning direct civil liability to an AI system would mainly benefit claimants by easing the burden of proof. The establishment of an e-personality could be accompanied by the assignment of a unique identification number to each AI system, along with the creation of a "register of artificial agents," modeled after traditional corporate registries. This would allow individuals interacting with intelligent agents to obtain information about their financial resources, ownership, and any harmful

theories, posing fundamental challenges to the legal frameworks designed to govern human conduct.[321]

Retribution, for example, is grounded in a commutative logic between the wrongful act and the corresponding punishment. However, artificial intelligence systems cannot be said to 'deserve' punishment, as they lack consciousness, intentionality, and *mens rea*.[322] Similarly, assigning responsibility to developers or users may fail to uphold the principles of culpability and causation.[323] In this sense, punishment — whether directed at the system itself or its 'creators' — risks offering a merely symbolic response to harm, falling short of addressing the systemic risks or design flaws embedded in the technology.[324] These shortcomings are often independent of any single operator and rather reflect, at a higher level, a *mala gestio* in the failure to adopt preventive measures concretely capable of avoiding harm.[325]

Within this framework, it appears more appropriate to conduct an analysis of the victim's needs and interests, shifting the focus away from blame and punishment toward reparation.

The general-preventive function of criminal sanctions — another cornerstone of criminal law — also proves problematic in the context of artificial intelligence.[326] It

---

incidents attributed to them—thus enabling better assessments of their reliability. On this topic see also Beck, S., *The Problem of Ascribing Legal Responsibility in the Case of Robotics*, AI & Society 31 (2016): 473 ss.

[321] Basile, F., *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, (2019): 31; Borsari, R., *Intelligenza artificiale e responsabilità penale: prime considerazioni*, *Medialaws* (2019): 262 et seq.

[322] Asaro, P.M., *A body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, in Lin, P., Abney, K., Bekey, G.A., *Robot Ethics*, (The MIT Press Cambridge, 2012), 169 et seq.; also ID., Determinism, Machine Agency, and Responsibility, *Politica & Società* (2014): 265 et seq., and ID., Robots and Responsibility from a Legal Perspective, *Proceedings of the 2007 IEEE International Conference on Robotics and Automation* (2011): 20 et seq.

[323] Asaro, P.M., *A body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, in Lin, P., Abney, K., Bekey, G.A., *Robot Ethics*, (The MIT Press Cambridge, 2012), 169 et seq.; also ID., Determinism, Machine Agency, and Responsibility, *Politica & Società* (2014): 265 et seq., and ID., Robots and Responsibility from a Legal Perspective, *Proceedings of the 2007 IEEE International Conference on Robotics and Automation* (2011): 20 et seq.

[324] Cappellini, A., *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 2019,

[325] Bartlett, B., *The possibility, op. cit.,* 269 et seq.

[326] Regarding so-called 'hard crimes', see the contribution of Nerantzi, E., Sartor, G. '*Hard AI Crimes': The Deterrence Turn*', in *Oxford Journal of legal studies* (2024); 44. According to the authors, the "AI deterrence paradigm" would serve to operationalize the "duty of care" that commissioners are expected to uphold when deploying AI agents with the technical features of economic machines. Specifically, distributors of AI agents would be required to maintain a compliance mechanism (when appropriate

presupposes the existence of conscious agents capable of understanding and adjusting their behaviours in response to legal sanctions, a condition that is not easily met in the case of autonomous systems.

However, the effectiveness of this principle is contingent on the capacity of the punished party to understand and adapt their behaviour. AI systems, being non-sentient entities, cannot learn from punishment in the way humans do. Moreover, punishing the organisations or individuals responsible for deploying AI systems may not effectively deter future misconduct if the root causes of harm lie in the inherent complexity and opacity of these technologies.

Rehabilitation, often considered the most forward-looking of the traditional theories of punishment, is equally inapplicable to AI.[327] The concept of rehabilitation presupposes the possibility of moral or psychological reform, a notion that is meaningless when applied to AI systems. While technical improvements can be made to AI systems, these do not constitute rehabilitation in the conventional sense. This raises questions about the utility and fairness of applying conventional punitive mechanisms to a fundamentally different type of agent.

The inadequacy of traditional punishment theories in addressing AI-related harm necessitates a re-examination of justice frameworks.[328] Philosophers[329] and legal

technologies are available in the relevant application domains), establishing: *machinae economicissimae*, which incorporate anticipated sanctions into their profit calculations; *machinae legales*, equipped with mandatory compliance systems; and *machinae legales et economicae*, which combine both approaches.

[327] Abbott R.B., Sarch, A.F., *Punishing Artificial Intelligence: Legal Fiction or Science Fiction*, in *UC Davis Law Review* (2019): 323.

[328] King, T.C., Aggarwal, N., Taddeo, M., Floridi, L., *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, in *Science and Engineering Ethics* (2019): 1 et seq.

[329] For a recent debate concerning the possible recognition of legal personhood for artificial intelligence systems, see Novelli, C., Floridi, L., and Sartor, G., *AI as Legal Persons: Past, Patterns, and Prospects*, in *Social Science Research Network* (2024): 1 et seq. The authors consider the attribution of legal personhood to AI to be unlikely, given that it would require substantial legislative reforms, which are not supported by any existing legal provisions. Past legal decisions have shown that legal personhood is typically conferred upon entities capable of fulfilling concrete social roles and holding rights and responsibilities, such as corporations. However, AI lacks characteristics such as the ability to own property or enter into contracts. Additionally, research indicates that the current legal trend does not support the recognition of AI as a legal entity. Future decisions will largely depend on legislative and political developments. Although the issue does not appear urgent in the short term, the ongoing evolution of AI technologies could make the discussion more relevant in the future, especially in contexts such as the integration of AI with human intelligence (e.g., through brain-machine interfaces). Nevertheless, for AI to be granted legal personhood, a major shift in the concept of legal personhood would be required, transforming AI from a separate entity into a potential extension of human identity.

scholars have argued for alternative approaches that shift the focus from punishing individual actors to addressing systemic issues and repairing harm.[330] This shift aligns with the principles of restorative justice, which prioritise accountability, dialogue, and the repair of relationships.

In this context, it is more appropriate to adopt an approach centered on the analysis of the victim's needs and interests, shifting the focus away from blame and punishment toward reparation.

The collective and systemic redefinition of accountability in the field of AI thus enables a more effective response to the challenges posed by artificial intelligence, taking into account the inherent unpredictability and opacity of such technological systems, and aiming both at the achievement of ethical standards and the reinforcement of public trust.[331]

Likewise, punishment must reflect the disconnection between the individual and culpability in the digital environment[332], thereby not only exposing the limitations of traditional punitive theories[333] but also encouraging the exploration of alternative responses to harm. These may include non-criminal and even strict liability mechanisms, which offer more effective and proportionate remedies beyond the boundaries of criminal law.[334]

1. Collective Risk Allocation in the AI Era

The deployment of AI systems requires a fundamental shift from individual liability models to collective risk allocation frameworks.[335] The development and

[330] Bartlett, B., *The possibility, op. cit.,* 269 et seq.

[331] Eichenhofer, J., *Trustworthy AI and Fundamental Rights*, in *European Review of Public Law* 36, no. 1 (2024): 131.

[332] Santoni de Sio, F., and G. Mecacci, G., *Four Responsibility Gaps with Artificial Intelligence: Why They Matter and How to Address Them*, *Philosophy & Technology* (2021): 34, identify four responsibility gaps, including the 'culpability gap'.

[333] Giannini, A., *Artificial intelligence, human oversight, and criminal liability: a european 'strenght tes*t', in *Criminalia*, (2021): 1 et seq.

[334] For a critical approach to strict liability, see Simons, K.W., *When Is Strict Criminal Liability Just?*, in *Journal of Criminal Law & Criminology*, vol. 87 (1997): 1075-76, discussing retributive views that denounce strict liability. Simons argues that 'strict liability appears to be a straightforward case of punishing the blameless, an approach that might have consequential benefits but is unfair on any retrospective theory of just deserts'. This highlights the concern that strict liability punishes individuals who are not at fault, which may be seen as unjust under a retributive justice framework.

[335] Mongillo, V., *Corporate criminal liability for AI-related crimes: possible techniques and obstacles*, in Picotti L. & Panattoni, B. (Eds.), *Traditional Criminal Law Categories and AI: Crisis or Palingenesis?*

implementation of AI involve multiple actors, including corporations, developers, policymakers, and regulators, whose decisions collectively shape the system's outcomes. Individual criminal liability, particularly in light of the complexity of algorithmic systems, often becomes problematic—especially with regard to establishing causal links. Likewise, it frequently proves ineffective in ensuring the effective protection of rights. This is particularly evident in high-risk contexts, where algorithmic outputs have significant consequences for society at large.

In this regard, accountability[336] represents the most appropriate solution for addressing harm caused by artificial intelligence, within a framework where collective risk allocation suggests that responsibility should be shared and distributed across the entire lifecycle of the technology, from its design to its deployment. What thus becomes essential is a thorough assessment of the actual risks posed by the tool, closely tied to the implementation of ongoing control and monitoring mechanisms, starting from the development phase and extending throughout the post-deployment stage.

Moreover, regulatory instruments should allow for the continuous adaptation of measures, in proportion to newly emerging risks.

This approach moves beyond the punitive logic of the traditional legal framework, reaffirming the criminal law principles of harm and ultima ratio. It seeks to embed ethical considerations into the fabric of AI governance, ensuring that risks are managed collectively rather than attributed to individual fault[337]. By framing responsibility as a 'shared endeavour', collective risk allocation aligns with broader societal interests in preventing harm and safeguarding fundamental rights. When

(Maklu, 2024), 82 et seq.; Consulich, F., *Criminal Law and Artificial Intelligence: Perspective From Italian And European Experience*, *European Criminal Law Review*, no. 3 (2023): 270-307.

[336] Accountability, in fact, is one of the seven principles mentioned by the Artificial Intelligence High-Level Expert Group (AI HLEG), established by the European Commission in the document *'Ethics Guidelines for Trustworthy AI'* and represents a recurring element in European regulatory initiatives, as well as in the broader global landscape of so-called AI ethics. For instance, reference can be made to the OECD principles (*OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, 2019*); the UNESCO recommendations on AI ethics (*UNESCO, Recommendation on the ethics of artificial intelligence, SHS/BIO/REC-AIETHICS/2021,* 2021). For a global overview, see the studies conducted by Fjeld, J. et al., *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI* (Berkman Klein Center for Internet & Security, 2020); Jobin, A. et al., *The global landscape of AI ethics guidelines*, in *Nature Machine Intelligence*, vol. 1 (2019).

[337] Rességuier, A., Rodrigues, R., *AI ethics should not remain toothless! A call to bring back the teeth of ethics*, in *Big Data & Society* (2020): 2.

regulatory or technical standards are not adhered to, an organisation can be held liable even in the absence of a direct causal link between its 'organisational fault' and the specific harmful event. This form of liability is ultimately linked to structural and systemic inadequacies, rather than to individual fault or a clear causal connection between action and outcome. From this perspective, organisational liability—particularly in cases of non-compliance with the standards established for a given category of risk—would take the form of strict liability.[338]

In the context of artificial intelligence, this calls for the prioritisation of proactive risk management, grounded in a logic of prevention[339]. Within this framework, accountability emerges as a key mechanism for fostering collective trust and safeguarding shared rights and values.

Moreover, the interdependence inherent in AI systems—whose performance relies heavily on the data used for training—highlights the importance of adopting an integrated approach not only to risk assessment but also to the distribution of responsibility. All stakeholders involved must contribute to effective risk mitigation.

Collaboration, particularly within governance structures, can thus play a crucial role in promoting a fair balance between inclusive development and transparency. This includes the establishment of mechanisms for stakeholder dialogue, data-sharing agreements, and ethical review boards, all aimed at ensuring that AI systems align with societal priorities and ethical standards.

2. Restorative Justice and the Sense of Sanction

Restorative justice offers a powerful lens through which to address the responsibility gap created by AI systems.[340] Unlike traditional punitive approaches, restorative justice focuses on understanding and repairing harm, addressing the needs of victims,

[338] Canato, M.C., *La responsabilità da reato degli enti di fronte al 'rischio da intelligenza artificiale': responsibility gap e 'rinnovata' corporate liability*, in *Quaderni del Dottorato in Giurisprudenza dell'Università di Padova 2024*, edited by Paola Lambrini (Ledizioni, 2025), 61-90.

[339] European Commission, *Communication on the Precautionary Principle*, COM(2000) 1 final, 2 February 2000.

[340] Ligi, T.K., *Introduction of restorative justice practices in criminal justice system: An overview'* in *International Journal of Criminal, Common and Statutory Law*, no. 4 (2024): 132-138; Nurul P. A. Nasution, Jubair and Abdul Wahid, *The Concept of Restorative Justice in Handling Crimes, European Journal of Law and Political Science* (2022).

and fostering accountability through collaborative processes.[341] In the context of AI, restorative justice could provide a framework for engaging all stakeholders—victims, developers, organisations, and regulators—in meaningful dialogue to address harm and prevent recurrence. Such a system not only enables a more tailored application of sanctions, but also serves as an effective instrument for identifying, through a proportional and dialogical approach, the most effective sanctions and the most appropriate remedial measures to address the needs of the harmed party, based on a case-by-case assessment.[342] In fact, victims of AI-induced errors often experience frustration and alienation due to the opaque and impersonal nature of these systems.[343] Restorative justice prioritises their involvement, ensuring that their voices are heard and their needs are addressed. For instance, victims might receive explanations of how errors occurred, assurances of corrective action, and commitments to systemic reform. This approach not only repairs trust but also reinforces the legitimacy of the justice system. By centring the experiences of those affected, restorative justice fosters a sense of empowerment and inclusion that is often

[341] The term 'accountability' derives from the Latin *computāre* (to calculate, consider), via Old French *acconter*, combined with the suffix *-ability*, denoting capacity. 'Responsibility,' by contrast, originates from *respondere* (to answer, to pledge). While 'liability' refers to specific legal obligations, accountability and responsibility operate within broader moral and social frameworks. Accountability may also be understood procedurally—as a public process of evaluating conduct to determine its justification and assign liability. This becomes especially significant in AI governance, where accountability mechanisms are essential to allocate risks and responsibilities among multiple stakeholders.

[342] The Principles of Restorative Practice outlined by the Restorative Justice Council (RJC) focus on core values that practitioners should uphold. These six principles are:

1. Restoration – The primary goal is to address and repair the harm caused.
2. Voluntarism – Participation in restorative processes is voluntary, with participants making an informed choice.
3. Neutrality – Restorative processes are fair and unbiased toward all participants.
4. Safety – The process ensures the safety of participants and provides a secure space for expressing feelings and views about the harm caused.
5. Accessibility – Restorative processes are non-discriminatory and accessible to everyone affected by conflict or harm.
6. Respect – Restorative practices uphold the dignity of all participants and those affected by the harm.

These principles guide restorative practices to ensure fairness, safety, and inclusivity.

[343] Users affected by AI-induced errors frequently report feelings of frustration and alienation, particularly due to the opaque and impersonal nature of algorithmic decision-making. See Reuben Binns et al., *It's Reducing a Human Being to a Percentage,* in *Perceptions of Justice in Algorithmic Decisions Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Association for Computing Machinery, 2018), https://doi.org/10.1145/3173574.3173951. The study highlights how lack of transparency and explainability in AI systems contributes to diminished user trust and a sense of powerlessness.

absent in traditional punitive processes.[344]

Restorative justice also invites reflection on the broader societal role of sanctions.[345] The primary aim of sanctions in this context should not be retribution but the effective protection of fundamental rights and the promotion of accountability.[346] Measures such as public apologies, commitments to improve AI transparency, and reparative actions, like funding initiatives to address biases in AI training data, can achieve these goals without resorting to criminal penalties. By shifting the focus away from punishment, restorative justice aligns with the principle of criminal law as an *extrema ratio*[347]reserved for the most severe breaches of societal norms. This shift also highlights the potential for restorative justice to serve as a bridge between legal and ethical accountability, ensuring that justice systems remain relevant and responsive in the age of AI.

Restorative processes can also address the psychological and societal impacts of AI-related harm. Victims, developers, and organisations can engage in structured dialogues to share perspectives, identify root causes, and co-develop solutions. These processes not only facilitate healing but also foster a culture of shared accountability and continuous improvement. By emphasising dialogue and collaboration, restorative justice can transform the responsibility gap into an opportunity for systemic learning

---

[344] Bertasini, I., '*Tra esigenze del giusto processo e prospettive rigenerative: il paradigma della Restorative Justice dentro e oltre il d.lgs. 150/2022'*, in *La scrittura nella giustizia. Una ricerca applicata di topica giudiziale*, (Padova University press, 2024), 243 et seq.

[345] Mannozzi, G., *Nuovi scenari per la giustizia riparativa. Riflessioni a partire dalla legge delega 134/2021*, in *Archivio Penale* (2022): 1. For a deeper understanding of the concept, see also Reggio, F., *Giustizia Dialogica. Luci e ombre della restorative justice* (Franco Angeli, 2010).

In her article, Mannozzi argues that restorative justice should enhance, not replace, traditional criminal law, making it more people-centered and community-focused. She highlights its potential to transform the justice system by prioritizing repair and community involvement. Federico Reggio, in his work on dialogic justice, explores both the benefits and challenges of restorative justice, emphasizing its role in modern legal systems. Both authors advocate for a justice approach that integrates accountability, community participation, and repair.

[346] Nonetheless, it is now evident that the traditional approach, focused on individual omissions or negligent behavior, must evolve to encompass the broader management of risks associated with the use of autonomous technologies. In this context, the AI Act serves as a critical benchmark, as it establishes clear requirements and provides essential guidelines to ensure organizations adopt continuous and proactive monitoring of their technologies. This reduces the potential 'responsibility gap' created by the decision-making autonomy inherent in AI systems.

[347] Florio*,* M.E., *Struggle in favour of a criminal law as an 'ultima ratio'. Critical observations on the criminalisation obligations arising from the jurisprudence of the European Court of Human Rights in the light of the principle of subsidiarity*, in *European Criminal Law Review,* no.13 (2023): 135-164.

and innovation.

3. Filling the Responsibility Gaps

The 'responsibility gap' in AI governance poses significant challenges to traditional legal systems. Restorative justice offers a practical and ethical response to these challenges by addressing harm in a holistic and inclusive manner.[348] For example, the dialogical tools typical of restorative justice could enable a deeper analysis of the root causes of erroneous AI outputs, allowing for a precise allocation of responsibility among the involved actors and the definition of targeted preventive measures. Moreover, within a socially shared framework, these mechanisms could function as spaces for information and awareness-raising, equipping stakeholders with the means to adapt risk categories and understand their practical implications—particularly in relation to distributed responsibility. The culture of accountability, therefore, intersects with restorative justice in fostering a collaborative governance model that upholds the principle of trust in the context of artificial intelligence.

Restorative justice also highlights the relevance of a forward-looking form of responsibility—one that prioritises prevention over repression—encouraging organisations to adopt ex ante risk management strategies. Additionally, dialogue with affected communities ensures the effectiveness of fundamental rights protection, within a long-term and adaptive framework.

Furthermore, restorative justice aligns with a collective vision of risk allocation by offering tailored solutions for residual harms that cannot be addressed through preventive measures alone and that thus fall within the scope of 'permitted risk.' If an AI system causes harm despite adherence to best practices during development and deployment, restorative mechanisms can be used to adapt responsibility to the specific needs of the harmed parties, ensuring both effective and proportionate redress.

This shared and integrated framework, in turn, enables a concrete modelling of 'legal risk'[349] arising from artificial intelligence, effectively re-aligning abstract technological risks with the principle of criminal *offensiveness*, thereby bridging the gap between

[348] Vidmar, N., and Miller, D.T., *Socialpsychological Processes Underlying Attitudes Toward Legal Punishment*, in *Law and Society Review* (1980): 565–602.

[349] Canato, M.C., *Verso il superamento del 'legal risk' europeo: intelligenza artificiale e approccio proporzionale al rischio*, *La Legislazione Penale* (2024).

potential harm and the thresholds required for penal intervention.

4. Conclusions

The 'responsibility gap'[350] in the context of artificial intelligence calls for a profound re-evaluation of traditional criminal law categories through the lens of accountability. Individual criminal liability—still anchored to the principles of culpability and the causal nexus between conduct and typical harm—proves increasingly inadequate in addressing the specific challenges posed by AI systems, particularly due to the multiplicity of actors involved in risk management. [351]

As a result, it becomes necessary to promote a shift towards collective risk allocation, grounded in a 'renewed' model of corporate liability and in the use of context-sensitive remedial tools, particularly through restorative justice frameworks.

Within this new paradigm, corporate liability is clearly decoupled from the prior model of individual criminal responsibility, which remains applicable only where personal culpability is clearly established. The analytical focus must therefore move away from the individual act and towards the systemic dimension of organisational accountability. This regulatory model also signals a gradual detachment from traditional criminal law, recognising that the specific risks posed by autonomous systems such as AI require a broader and more proactive approach. Failure to comply with the standards set by the AI Act and sector-specific legislation may thus give rise to strict liability, at least on a presumptive basis, with respect to AI-induced harm—even in the absence of a proven causal link or individual fault.

In this framework, prevention, transparency, and proactive risk management assume a central role, shifting the focus away from culpability as the main foundation of liability. Individual responsibility, where it arises, must be reinterpreted within a systemic and distributed logic. At the same time, corporate liability becomes autonomous from that of its members, aligning organisational duties with the need to manage systemic risks

[350] Matthias, A., *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, *op. cit.*

[351] Consulich, F., *Il nastro di Möbius. Intelligenza artificiale e imputazione penale nelle nuove forme di abuso del mercato*, in *Banca Borsa Titoli di Credito* (Giuffrè, 2018), 195 et seq.

inherent to AI. [352]

In this context, restorative justice provides an opportunity to explore sanctioning mechanisms and adaptive remedies that are effectively concrete in managing AI-related harm. It promotes a model that shifts attention from punitive logic to prevention, deterrence, and repair, in accordance with the general and special preventive functions of penalties.[353]

In the AI era, the legal framework evolves ethically, acknowledging that when faced with 'accepted' and 'collectively' borne risks, the protection of fundamental human rights must serve as the primary benchmark for any form of justice. The move away from criminal law—consistent with its nature as an *extrema ratio*—and the promotion of dialogical and individualised instruments, such as restorative justice, ultimately allows for the reconstruction of a truly 'trustworthy European AI'.[354]

Adopting these innovative frameworks, including a renewed corporate liability and restorative justice, will be crucial in ensuring that justice remains both effective and equitable, especially in the face of rapid technological advancements. Addressing the unavoidable risks of artificial intelligence lies in conducting thorough and proportionate risk assessments[355], ensuring a careful human balance between costs and benefits, maintaining effective human oversight, and fostering a collective approach to risk management through robust accountability mechanisms.[356]

In conclusion, when considering the risk-based approach integral to the AI Act, the criteria for determining liability must undergo significant transformation to address the inherently unavoidable risks posed by AI systems. The current paradigms of criminal law, particularly those grounded in the strict application of principles like legality and

[352] The AI Act treats artificial intelligence as a 'product'. In this context, the regulation aligns with frameworks that impose strict liability for harmful events, assigning responsibility to the owner or operator of the system, even when it is intended for public or collective use.
This approach underscores the principle that liability can arise independently of direct fault, emphasizing accountability for managing the risks inherent in the development, deployment, and oversight of such technologies.

[353] Salvi, F., La funzione della pena tra castigo e risocializzazione, *SalvisJuribus* (2022).

[354] European Commission, *Ethics Guidelines for Trustworthy AI*, High-Level Expert Group on Artificial Intelligence, April 2019.

[355] Novelli, C., Casolari, F., Rotolo, A., Taddeo, M., Floridi, L., *AI Risk Assessment, op. cit.*

[356] Canato, M.C., *Verso il superamento del 'legal risk' europeo: intelligenza artificiale e approccio proporzionale al rischio, op. cit*.

culpability, must be reassessed and, in some cases, replaced with alternative protective frameworks. These frameworks should ensure that sanctions for violations of fundamental rights remain 'proportionate and effective', in line with the evolving principles articulated by supranational jurisprudence[357]. Only through this recalibration, incorporating strict 'non-criminal' corporate liability and restorative justice, can justice systems effectively address the challenges posed by AI. This approach ensures the preservation of human oversight. It maintains a human-centred framework for criminal law, applying the principle of criminal liability as an *extrema ratio*, and only when negligence is clearly evident.

[357] The Strasbourg case law, with regard to medical liability, has ruled out the requirement for criminal sanctions by States in cases of medical negligence, as long as the effective protection of the right to life under Article 8 of the European Convention on Human Rights (ECHR) is guaranteed. For a detailed examination of this issue, see Pranka, D., *The Price of Medical Negligence – Should it be Judged by the Criminal Court in the Context of the Jurisprudence of the European Court of Human Rights?,* in *Baltic Journal of Law & Politics,* no. 14, 1 (2021); 124-152.

Bibliography

Abbott, R. B., Sarch A. F., 'Punishing Artificial Intelligence: Legal Fiction or Science Fiction', *UC Davis Law Review* (2019).

Asaro, P. M., 'Robots and responsibility from a legal perspective', *Proceedings of the 2007 IEEE International Conference on Robotics and Automation* (2011).

Asaro, P. M., 'A Body to Kick, but still no soul to damn: Legal perspectives on robotics', in *Robot Ethics* edited by Lin P., Abney K., Bekey G. A. (The MIT Press Cambridge, 2012),

Asaro, P. M., 'Determinism, machine Agency, and responsibility', *Politica & Società* (2014).

Bartlett, B., 'The Possibility of AI-Induced Medical Manslaughter: Unexplainable Decisions, Epistemic Vices, and a New Dimension of Moral Luck', *Medical Law International* (2023).

Basile, F., 'Intelligenza Artificiale e Diritto Penale: Quattro Possibili Percorsi di Indagine', *Diritto Penale e Uomo* (2019).

Beck, S. 'The Problem of Ascribing Legal Responsibility in the Case of Robotics', *AI & Society* 31 (2016): 473 ss.

Bertasini, I., 'Tra esigenze del giusto processo e prospettive rigenerative: il paradigma della restorative justice dentro e oltre il D.Lgs. 150/2022', *La scrittura nella giustizia. Una ricerca applicata di topica giudiziale (*2024).

Canato, M. C., 'Verso il Superamento del 'Legal Risk' Europeo: Intelligenza Artificiale e Approccio Proporzionale al Rischio', *La Legislazione Penale* (2024).

Canato, M. C., 'La Responsabilità da Reato degli Enti di Fronte al 'Rischio da Intelligenza Artificiale': Responsibility Gap e 'Rinnovata' Corporate Liability', in *Quaderni del Dottorato in Giurisprudenza dell'Università di Padova 2024*, edited by P. Lambrini (Ledizioni, 2025).

Cappellini, A., 'Machina delinquere non potest? Brevi Appunti su Intelligenza Artificiale e Responsabilità Penale', *Criminalia* (2019).

Consulich, F., 'Rischio consentito', in *Reato colposo*, *Enciclopedia del Diritto*, ed. M. Donini (Milano: Giuffrè, 2021).

Consulich, F., 'Il concorso di persone nel reato colposo' (Torino: 2023).

Consulich, F., 'Criminal Law and Artificial Intelligence: Perspective From Italian and European Experience', *European Criminal Law Review*, no.3 (2023).

Consulich, F., 'Il Nastro di Möbius. Intelligenza Artificiale e Imputazione Penale nelle Nuove Forme di Abuso del Mercato', *Banca Borsa Titoli di Credito*, (Giuffrè, 2018).

Consulich, F., 'Rischio Consentito' in *Enciclopedia del diritto. Reato colposo*, edited by M. Donini (Giuffrè, 2021).

Darshan, V., 'Demystifying the Role of Artificial Intelligence in Legal Practice', *Nirma University Law Journal*, no. 2 (2019).

Doncieux, S., J. B. Mouret, 'Beyond Black-Box Optimization: A Review of Selective Pressures for Evolutionary Robotics' , *Evolutionary Intelligence* (2014).

Floridi, M. E., 'Struggle in Favour of a Criminal Law as an 'Ultima Ratio.' Critical Observations on the Criminalisation Obligations Arising from the Jurisprudence of the European Court of Human Rights in the Light of the Principle of Subsidiarity', *European Criminal Law Review*, no. 2 (2023).

Eichenhofer, J., 'Trustworthy AI and Fundamental Rights', *European Review of Public Law* 36, no. 1 (2024): 131.

Fragasso, B., 'La Responsabilità Penale del Produttore di Sistemi di Intelligenza Artificiale.' *Sistema Penale* (2023).

Fragasso, B., 'Intelligenza artificiale e responsabilità penale: principi e categorie alla prova di una tecnologia "imprevedibile*"*'. (Turin: Giappichelli, 2025).

Fjeld, J., et al., 'Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI' (Berkman Klein Center for Internet & Security, 2020).

Forti, G., 'Colpa ed Evento nel Diritto Penale' (Giuffrè, 1990).

Giannini, A., 'Artificial Intelligence, Human Oversight, and Criminal Liability: A European 'Strength Test'', *Criminalia* (2021).

Giannini, A. 'Criminal Behavior and Accountability of Artificial Intelligence Systems' (Maastricht Law Series, 2023).

Gless, S., E. Silverman, T. Weigend, 'If Robots Cause Harm, Who Is to Blame?

Self-Driving Cars and Criminal Liability', *New Criminal Law Review* (2016).

Hallevy, G., 'The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control', *Akron Intellectual Property Journal* (2010).

Hallevy, G., 'Liability for Crimes Involving Artificial Intelligence Systems' (Springer, 2015).

Hilgendorf, E., 'Automated Driving and the Law.' In *Robotics, Autonomics and the Law*, edited by E. Hilgendorf and U. Seidel (Nomos, 2017).

Jobin, A., et al., 'The Global Landscape of AI Ethics Guidelines', *Nature Machine Intelligence*, no. 1, (2019).

King, T. C., Aggarwal N., Taddeo M., and Floridi L., 'Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions', *Science and Engineering Ethics*, no. 2 (2019).

Lima, D., 'Could AI Agents Be Held Criminally Liable: Artificial Intelligence and the Challenges for Criminal Law' *South Carolina Law Review* (2018).

Mannozzi, G., 'Nuovi Scenari per la Giustizia Riparativa. Riflessioni a Partire dalla Legge Delega 134/2021', *Archivio Penale* (2022).

Mongillo, V., 'Corporate Criminal Liability for AI-Related Crimes: Possible Techniques and Obstacles', in *Traditional Criminal Law Categories and AI: Crisis or Palingenesis?*, edited by Picotti L. and Panattoni B., (Maklu, 2023).

Nerantzi, E., Sartor G., ''Hard AI Crimes': The Deterrence Turn' (Oxford, 2024).

Novelli, C., 'L'Artificial Intelligence Act Europeo: Alcune Questioni di Implementazione', *federalismi.it*, no.2 (2024).

Novelli, C., Casolari F., Rotolo A., Taddeo M., Floridi L. 'AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act', *Digital Society* (2024).

Novelli, C., Floridi L., Sartor G., 'AI as Legal Persons: Past, Patterns, and Prospects.' *Social Science Research Network* (2024).

Ortalda, A., De Hert P., 'Artificial Human Rights Impact Assessment.' In *Artificial Intelligence and Human Rights*, edited by A. Quintavalla and J. Temperman (Hart Publishing, 2023).

Pagallo, U., 'The Laws of Robots: Crimes, Contracts and Torts' (Springer, 2013).

Panattoni, B., 'AI and Criminal Law: the Myth of 'Control' in a Data-Driven Society', *Revue Internationale de Droit Pénal* (2021).

Pranka, D., 'The Price of Medical Negligence – Should It Be Judged by the Criminal Court in the Context of the Jurisprudence of the European Court of Human Rights*?'* (Kaunas, 2021).

Reggio, F., 'Giustizia Dialogica: Luci e Ombre della Restorative Justice' (Franco Angeli, 2010).

Rességuier, A., Rodrigues R., 'AI Ethics Should Not Remain Toothless! A Call to Bring Back the Teeth of Ethics', *Big Data & Society* (2020).

Roxin, C., Greco L., 'Strafrecht, Allgemeiner Teil', (C.H. Beck, 2020).

Salvadori I., 'Agenti Artificiali, Opacità Tecnologica e Distribuzione della Responsabilità Penale', *RIDPP* (2021).

Salvi, F., 'La funzione della pena tra castigo e risocializzazione', *Salvis Juribus* (2022).

Santoni de Sio, F., Mecacci G., 'Four Responsibility Gaps with Artificial Intelligence: Why They Matter and How to Address Them.' *Philosophy & Technology* (2021).

Simons, K. W., 'When Is Strict Criminal Liability Just?' *Journal of Criminal Law & Criminology* (1997).

Surden, H., M. A. Williams, 'Technological Opacity, Predictability, and Self-Driving Cars', *Cardozo Law Review* (2016).

Taylor, I., 'Collective Responsibility and Artificial Intelligence', *Philos. Technol.* Vol. 37, n. 27 (2024).

Van de Poel, I., Sand, M., 'Varieties of responsibility: two problems of responsible innovation', *Synthese* 198 (Suppl 19), 4769–4787 (2021).

Vidmar, N., Miller D. T., 'Socialpsychological Processes Underlying Attitudes Toward Legal Punishment.', *Law and Society Review* (1980).