REVIEW OF HUMAN RIGHTS
Vol. 12, No. 1, Winter 2026, pp: 45-59
DOI: 10.35994/rhr.v12i1.318

# Artificial Intelligence and the Transformation of Human Rights: From Justification to Probability

Marco Marsili•
https://orcid.org/0000-0003-1848-9775

**Abstract**
Artificial intelligence is not merely a set of technical tools; it acts as a systemic force reshaping the architecture of human rights protection. This article argues that AI introduces a structural shift from traditional decision-making based on justification and human accountability to probabilistic, opaque, and adaptive governance. Drawing on critical theory, legal scholarship, and selected empirical examples from law enforcement, social media governance, and welfare systems, the study demonstrates how AI diffuses responsibility, constrains remedies, and redefines states' positive obligations. The analysis situates AI as a systemic actor that challenges the foundational assumptions of international human rights law, emphasizing the cumulative and transversal nature of its impact.

**Keywords**: Artificial Intelligence, Human Rights, Responsibility, Positive Obligations, Remedies, Systemic Risks, Probabilistic Governance.

## Introduction

The rapid integration of artificial intelligence into public and private decision-making has profoundly reshaped the landscape of human rights protection. Over the past decade, international organisations, courts, and scholars have increasingly focused on the risks posed by algorithmic systems to specific rights, including privacy, non-discrimination, freedom of expression, and due process. This growing body of work has significantly advanced the normative and legal understanding of artificial intelligence, producing an extensive set of principles, guidelines, and policy recommendations aimed at ensuring human-centric and rights-respecting technological development.[1]

• Marco Marsili, Cà Foscari University of Venice, is an associate researcher at the Centre for International Studies (CEI–Iscte), Iscte – University Institute of Lisbon, Portugal. He holds a European Ph.D. in History, Studies of Security and Defense, and an advanced post-graduation in Armed Conflicts and Human Rights. Email: marco_marsili@iscte.pt

Published Online: September 20, 2026.
ISSN (Print): 2520-7024; ISSN (Online): 2520-7032.
https://reviewhumanrights.com

Despite this progress, much of the existing literature remains oriented toward identifying discrete violations or sector-specific risks, often treating artificial intelligence as a new regulatory object rather than as a structural factor capable of transforming the conditions under which human rights operate. In many cases, AI-related challenges are framed as compliance problems to be addressed through transparency obligations, impact assessments, or risk mitigation strategies. While these approaches are not without value, they tend to assume that the underlying architecture of international human rights law remains conceptually stable. This assumption, however, has increasingly been called into question.

Artificial intelligence systems do not merely introduce new tools for decision-making; they alter the epistemic and institutional foundations on which legal responsibility, justification, and accountability are built. Automated and semi-automated systems rely on probabilistic reasoning, opacity, distributed agency, and anticipatory risk management, features that sit uneasily with a legal framework historically grounded in ex post adjudication, identifiable duty-bearers, and reason-giving decisions.[2] Scholars working at the intersection of law and technology have noted that these characteristics challenge core legal concepts such as causation, foreseeability, and attribution, which are central to the operation of human rights law.[3]

Importantly, this tension does not necessarily manifest itself in explicit or easily justiciable violations. Rather, it operates at a systemic level, gradually eroding the effectiveness of human rights guarantees while leaving formal legal commitments largely untouched. As the Office of the High Commissioner for Human Rights (OHCHR) has repeatedly acknowledged, many of the most significant risks associated with artificial intelligence arise not from isolated abuses, but from the cumulative and structural effects of data-driven and automated decision-making on governance and accountability.

This article starts from the premise that artificial intelligence should therefore be analysed not only as a source of individual rights infringements, but as a driver of structural transformation in the architecture of human rights protection. In this sense, artificial intelligence constitutes a systemic risk: not because it inevitably violates rights, but because it reshapes the legal conditions under which rights are defined, attributed, and enforced. The risk lies in the displacement of a paradigm centred on justiciable claims, responsibility, and justification by a model of probabilistic governance focused on prediction, optimisation, and risk management.

Recent contributions have begun to gesture toward this broader transformation, particularly in discussions of algorithmic governance and accountability. Authors have highlighted the emergence of responsibility gaps, the growing reliance on automated scoring and classification, and the difficulty of ensuring meaningful human oversight in complex socio-technical systems. An expert report prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence identifies systemic risks, concentrations of digital power, and threats to procedural legitimacy

arising from AI, thereby extending the analysis beyond infringements of individual rights.[4]

Yet these analyses often remain fragmented or confined to specific regulatory domains. Legal scholarship tends to focus on the application of existing human rights standards to particular uses of artificial intelligence, while critical and interdisciplinary approaches frequently diagnose deeper transformations without fully engaging with the internal logic and constraints of international human rights law. As a result, the systemic implications of artificial intelligence for the human rights framework as a whole remain under-theorised.

Against this background, the article addresses the following research question: how does the increasing reliance on artificial intelligence systems transform the structural conditions under which human rights are defined, attributed, and enforced in international human rights law? The central claim advanced here is that artificial intelligence introduces a shift from ex post justiciable rights grounded in responsibility and reason-giving toward ex ante probabilistic outputs governed through risk management. This shift challenges core pillars of the human rights framework, including the attribution of responsibility, the scope of positive obligations, and the practical accessibility of effective remedies under instruments such as the International Covenant on Civil and Political Rights (ICCPR)[5] and the European Convention on Human Rights (ECHR).[6]

The article contributes to the existing literature by offering a systemic and architecture-level analysis of artificial intelligence and human rights, rather than a catalogue of sectoral risks. In doing so, it builds upon and extends recent work on technologically mediated power, hybrid forms of governance, and the legal implications of emerging technologies for normativity and agency. While these earlier analyses focused on security and conflict-related environments, the present article demonstrates that similar dynamics of diffusion of responsibility, opacity, and anticipatory control are increasingly embedded in civilian governance contexts shaped by artificial intelligence.

Methodologically, the article adopts a theoretical–critical legal analysis grounded in international human rights law. It draws on primary normative sources, including international treaties and authoritative interpretative instruments, as well as selected doctrinal and interdisciplinary scholarship, to reconstruct the conceptual assumptions underlying human rights protection. Rather than proposing immediate regulatory solutions, the analysis aims to clarify the nature of the systemic risk posed by artificial intelligence, thereby providing a more solid foundation for future legal and institutional responses.

## 1. Methodology

This article adopts a theoretical–critical legal methodology grounded in international human rights law. The methodological choice reflects the nature of the research question, which is normative and structural in character and does not

concern the empirical measurement of artificial intelligence impacts or the evaluation of specific regulatory instruments. Rather, the analysis focuses on how AI-mediated decision-making transforms the structural conditions under which human rights are conceptualised, attributed, and enforced.

The analysis is anchored in primary sources of international human rights law, including binding treaties such as the ICCPR and the ECHR, as well as authoritative interpretative materials produced by international and regional bodies. These sources are approached through a systemic method of interpretation, consistent with established human rights adjudicatory practice, which treats legal norms as elements of an interdependent legal architecture rather than as isolated provisions.[7] Particular attention is devoted to concepts such as responsibility, justification, foreseeability, and access to remedies, understood as structural preconditions for the effectiveness of human rights protection.

In parallel, the article engages selectively with doctrinal and interdisciplinary scholarship on artificial intelligence, algorithmic governance, and law. Rather than offering an exhaustive review, it draws on influential contributions that explicitly address the interaction between technological mediation and legal normativity. In this context, scholarship emphasising the role of law as a framework for reason-giving and contestability in computational environments is particularly relevant,[8] as are analyses highlighting the normative implications of governance models based on automation and risk management.[9]

A central methodological choice concerns the level of abstraction at which artificial intelligence is examined. The article deliberately refrains from focusing on specific technical architectures or isolated use cases. Instead, artificial intelligence is treated as a socio-technical mode of governance characterised by probabilistic reasoning, opacity, distributed agency, and anticipatory control. This abstraction is methodologically justified insofar as the aim is to capture systemic effects that cut across sectors and legal domains, a move that has been endorsed in critical analyses of algorithmic governance and emerging technologies.[10] Concrete examples are invoked only where necessary to illustrate broader structural dynamics, without collapsing the analysis into case-based assessment.

The critical dimension of the methodology consists in exposing the implicit premises underpinning both regulatory responses and scholarly analyses of artificial intelligence and human rights. In particular, the article interrogates the assumption that existing human rights frameworks can absorb AI-related challenges through incremental adaptation, such as enhanced transparency or risk mitigation obligations. By contrast, the methodological stance adopted here treats artificial intelligence as a factor capable of altering the grammar of human rights law itself, thus requiring a reassessment of foundational concepts rather than their mere extension.

This approach is informed by prior theoretical work on technologically mediated forms of power and normativity, including analyses of hybrid and

cognitive environments in which agency and responsibility are diffused rather than directly exercised.[11] While the present article does not address security or conflict settings, it draws on similar methodological intuitions, namely that technological systems may generate legal effects indirectly by reshaping decision-making environments. The methodology thus remains firmly legal in orientation, while remaining open to insights from critical theory and socio-legal studies.

Finally, the article distinguishes its objectives from policy-oriented or solution-driven scholarship. It does not seek to propose immediate regulatory reforms or technical safeguards. Instead, it aims to clarify the nature of the systemic risk posed by artificial intelligence to the human rights framework, understood as a coherent legal architecture rather than a collection of isolated entitlements. This clarification constitutes a necessary preliminary step for any meaningful normative or institutional response.

## 2. State of the Art

The relationship between artificial intelligence and human rights has generated an extensive and rapidly expanding body of scholarship over the last decade. Early contributions were largely framed in ethical terms, focusing on values such as fairness, transparency, and accountability. While these approaches played a significant role in shaping public and institutional awareness, they were often detached from the structures of international human rights law, relying on moral principles rather than legally grounded obligations.[12]

Subsequent scholarship, particularly from the mid-2010s onward, increasingly adopted a legal and human rights–based perspective. International organisations have played a central role in this shift. The United Nations, through the Office of the High Commissioner for Human Rights, has repeatedly emphasised that artificial intelligence may interfere with a wide range of protected rights and has called for the application of existing human rights norms to AI-enabled systems. Similarly, within the European human rights system, concerns have been raised regarding the compatibility of automated decision-making with fundamental principles such as legality, proportionality, and effective judicial protection.[13]

In doctrinal legal scholarship, much of the analysis has followed a rights-by-rights approach. Authors have examined how algorithmic systems may affect specific entitlements, including privacy, equality, freedom of expression, and procedural fairness. This literature has clarified important doctrinal issues, such as the relevance of discrimination law to algorithmic bias,[14] the applicability of data protection and privacy standards to machine learning systems,[15] and the role of human oversight in automated decision-making processes. These contributions have strengthened the legal framing of AI-related harms and have supported the integration of human rights considerations into regulatory debates.

At the same time, an increasing number of scholars have pointed to the limits of this sectoral approach. Several analyses argue that artificial intelligence challenges foundational legal concepts, including agency, causation, and accountability, which

cannot be adequately addressed through incremental doctrinal adjustments alone. Particular attention has been devoted to responsibility gaps and to the opacity of machine learning systems, which complicate the attribution of legal responsibility and undermine traditional models of accountability.[16]

Despite these advances, the literature remains fragmented along disciplinary and analytical lines. Legal analyses often prioritise compliance mechanisms and enforcement tools, while interdisciplinary and critical scholarship tends to diagnose broader transformations in governance and power without systematically engaging with the internal logic of international human rights law. As a result, the cumulative and structural effects of artificial intelligence on the human rights framework as a whole remain insufficiently theorised.

A further characteristic of the current state of the art is the growing reliance on risk-based regulatory paradigms. Risk assessment, classification, and mitigation have become organising principles of AI governance, including within human rights–oriented regulatory initiatives. While this approach offers pragmatic advantages, it also reflects an implicit shift toward probabilistic reasoning as a basis for legal and institutional decision-making. The compatibility of such a shift with a human rights system grounded in individual claims, ex post adjudication, and effective remedies has received comparatively limited attention.[17]

Some recent theoretical contributions have begun to explore these tensions more explicitly, questioning whether human rights law is conceptually equipped to address technologies that operate through prediction, anticipation, and optimisation rather than discrete decision points. These analyses suggest that the most significant implications of artificial intelligence may lie not in overt norm violations, but in the gradual reconfiguration of the conditions under which legal norms operate. However, they often stop short of reconstructing the implications of this transformation for the architecture of international human rights law itself.

Within this evolving debate, prior research on technologically mediated power and hybrid forms of governance has highlighted how legal normativity may be indirectly affected by shifts in agency and decision-making environments. These insights, developed in adjacent fields, point toward the need for a systemic account of artificial intelligence and human rights that moves beyond sectoral analysis.

Against this background, the present article positions itself at the intersection of legal doctrine and critical theory. Rather than cataloguing AI-related risks to individual rights, it seeks to articulate a coherent account of artificial intelligence as a systemic risk to the architecture of human rights protection. By focusing on the shift from justificatory, responsibility-based decision-making toward probabilistic governance, the article aims to fill a conceptual gap in the existing literature and to provide a framework capable of explaining why many AI-related challenges resist resolution through conventional human rights tools.

## 3. Artificial Intelligence and the Shift from Justification to Probability

Artificial intelligence reshapes the conditions under which human rights operate not simply by producing potential violations, but by transforming the epistemic and normative foundations of law. Traditional human rights frameworks rely on ex post reasoning: they assume that decisions are made by identifiable agents who are accountable for their choices, that harms can be traced and justified, and that remedies can be administered following violation. Justification, reason-giving, and attribution form the conceptual pillars on which responsibility and effective protection rest.

AI-mediated decision-making, by contrast, introduces a mode of governance oriented toward prediction, optimisation, and probabilistic assessment.[18] Decisions are generated by systems that aggregate large datasets, identify correlations often beyond human comprehension, and produce outputs that are inherently opaque. Responsibility is diffused, causal chains are non-linear, and the justificatory structures expected by human rights law are weakened. Probabilistic outputs replace reasoned decisions; anticipatory risk management substitutes for ex post accountability. This transformation does not necessarily produce identifiable violations, but it systematically reshapes the conditions for the attribution of legal and moral responsibility.[19]

The shift from justification to probability manifests along multiple dimensions. First, it challenges the attribution of responsibility. Traditional law presumes that responsibility can be traced to human agents; AI introduces opacity, adaptivity, and probabilistic reasoning that make causal and normative attribution difficult. Automated systems often act as intermediaries, mediating or even co-producing decisions, such that nominally responsible actors cannot fully justify or explain outcomes. Existing approaches, such as delegated accountability or supervisory responsibility, address symptoms but do not resolve the structural transformation of agency.[20] Responsibility may need to be reconceptualised as distributed, networked, or layered, without undermining enforceability.[21]

Second, the shift affects positive obligations. States are required to take proactive steps to safeguard rights. Yet AI systems operate across jurisdictions and sectors, generating consequences that may not be attributable to any single agent. Discrete compliance measures, such as bias mitigation, transparency mandates, or auditing requirements, fail to address the systemic diffusion of influence. Addressing positive obligations under AI requires anticipatory frameworks that monitor structural effects, including periodic systemic audits, algorithmic impact assessments, and governance processes designed to detect emergent risks before they crystallise into individual harms.[22]

Third, the accessibility of remedies is affected. Ex post adjudication presupposes that harms can be traced to identifiable acts or omissions and that affected individuals can challenge these acts meaningfully. AI-mediated harms are often cumulative, opaque, or probabilistic. For instance, in predictive policing systems, risk scores generated from historical crime data can disproportionately

affect particular communities without a single decision-maker being responsible, making traditional remedies ineffective. In welfare administration, automated eligibility determinations based on algorithmic scoring may deny access to benefits, yet the logic behind such decisions may be inaccessible even to the administering agency. These examples illustrate how the procedural accessibility of rights is eroded, even when formal legal guarantees remain intact.

The implications extend beyond the individual rights level, affecting the broader architecture of human rights protection. Responsibility, positive obligations, and remedies are mutually reinforcing: weakening one pillar affects the others. As AI systems diffuse accountability, limit the scope of proactive obligations, and reduce remedy effectiveness, the systemic coherence of human rights law is undermined. This systemic perspective explains why sectoral analyses often underestimate AI's implications: the most significant effects are indirect, cumulative, and structural.

Concrete illustrations further clarify this systemic shift. In law enforcement, predictive policing algorithms rely on probabilistic patterns derived from historical data. Even if the algorithms are technically unbiased, they can perpetuate structural inequalities by embedding historical discrimination into automated decision-making. Responsibility is difficult to assign, positive obligations are diffuse, and remedies for affected individuals are largely inaccessible. Similarly, in social media governance, automated content moderation systems operate on massive scales, using machine learning to flag content probabilistically. Misclassifications or bias in these systems may have wide-ranging effects on freedom of expression, yet accountability is dispersed across platform administrators, engineers, and algorithmic processes. Welfare systems increasingly employ automated eligibility determination; opaque scoring systems may exclude individuals without identifiable human oversight. These examples, spanning criminal justice, online expression, and social services, demonstrate that the shift from justification to probability operates across domains, affecting both public and private actors.

The shift also challenges legal doctrine at a conceptual level. Human rights law assumes that rights violations can be attributed, contested, and remedied; AI complicates all three. Legal accountability is based on the ability to trace harm to a responsible agent; AI obscures this chain. Positive obligations presuppose that states can implement corrective or preventive measures; AI diffuses responsibility and introduces unpredictability. Remedies presuppose that legal and procedural mechanisms allow affected individuals to vindicate rights; AI-generated harms may be opaque or probabilistic, rendering remedies ineffective. The cumulative effect is a structural tension between the formal guarantees of law and the operational logic of AI.[23]

A critical implication of this analysis is that legal frameworks may need to reconceptualise key categories of normativity. Responsibility might be understood as distributed rather than individual; remedies may require proactive auditing rather than reactive adjudication; justification may involve transparency of process rather

than explanation of discrete outcomes. Such reconceptualisation does not abandon human rights law but demands a reinterpretation of its core pillars in light of technological mediation. By foregrounding these structural shifts, the article positions AI not as a collection of discrete threats, but as a systemic actor reshaping the architecture of human rights protection. Recognising this transformation is a necessary prerequisite for meaningful legal, institutional, and ethical responses.

Finally, this systemic perspective invites reconsideration of human rights scholarship itself. Traditional doctrinal analysis, focused on discrete rights and isolated harms, risks underestimating the indirect and structural effects of AI. Conversely, interdisciplinary and critical perspectives often highlight systemic transformations but fail to anchor analysis in legally enforceable norms. By integrating doctrinal, institutional, and critical insights, a more robust understanding of AI as a systemic driver of change emerges. This shift is normatively significant because international human rights law is structurally premised on contestability, reason-giving, and the possibility of ex post review, all of which are strained by probabilistic and anticipatory forms of governance. Rather than offering an exhaustive review, this article deliberately focuses on contributions that illuminate structural transformations in responsibility, accountability, and remedies. As summarized in Table 1, the shift from justification to probability affects all core pillars of human rights protection.

| Concept | Traditional Human Rights Framework | AI-Mediated Transformation |
|---|---|---|
| **Responsibility** | Individual, traceable to human decision-making | Distributed, networked, partially opaque |
| **Positive obligations** | Proactive duties by states, identifiable agents | Structural, anticipatory, multi-actor |
| **Remedies** | Ex post, legal recourse for identifiable harms | Proactive auditing, process transparency, limited ex post effectiveness |
| **Decision logic** | Justification-based (reasoned decisions) | Probability-based (predictive, adaptive outputs) |
| **Accountability** | Direct and attributable | Diffused across systems and actors |

*Table 1. Transformation of Core Human Rights Concepts under AI*

## 4. Implications for Responsibility, Positive Obligations, and Effective Remedies

The systemic transformation introduced by artificial intelligence has immediate implications for the three pillars of human rights law: the attribution of responsibility, the fulfillment of positive obligations, and the accessibility of effective remedies. These dimensions, long treated as mutually reinforcing, are increasingly strained under the conditions of AI-mediated governance.

Responsibility, traditionally conceived as the capacity of identifiable agents to justify and answer for their actions, is complicated by the diffusion of decision-making inherent in AI systems. Automated processes often operate as intermediaries between humans, or even as semi-autonomous actors in hybrid human-machine arrangements. Legal doctrine presumes that responsibility can be traced along causal chains; AI introduces opacity, adaptivity, and probabilistic reasoning that make causal and normative attribution difficult. Existing approaches to responsibility gaps—through principles such as delegated accountability or the attribution of outcomes to supervising authorities—offer partial solutions but fail to address the deeper structural reconfiguration of agency.[24] The implication is that human rights law may need to reconceptualize responsibility as distributed, networked, or layered, rather than exclusively individual, without undermining enforceability.

Positive obligations, which require states to take proactive measures to safeguard rights, are similarly challenged. AI systems operate across sectors and jurisdictions, generating outcomes that may be harmful even when no specific agent is at fault. Traditional regulatory instruments, including legislation, oversight mechanisms, or auditing requirements, tend to address discrete risks rather than the systemic diffusion of influence and control. Consequently, fulfilling positive obligations demands a shift from reactive interventions to anticipatory frameworks capable of monitoring structural dynamics. This may involve periodic systemic audits, the design of institutional checks that account for AI's emergent behaviors, or the integration of algorithmic impact assessments directly into governance processes.

The accessibility of effective remedies is equally affected. Ex post adjudication presupposes that harms can be traced to identifiable acts or omissions and that affected individuals can challenge these acts meaningfully. In AI-mediated contexts, harms may be cumulative, opaque, or probabilistic, rendering traditional litigation and complaint mechanisms insufficient.[25] Remedies may require proactive disclosure of algorithmic processes, systemic transparency measures, and procedural innovations enabling individuals to contest decisions whose logic is inaccessible to human understanding. Without such adaptations, rights may remain formally recognised but practically unenforceable, eroding the legitimacy of the human rights framework.

Together, these transformations reveal a tension between the formal structures of human rights law and the operational logic of AI systems. They suggest that conventional doctrines of responsibility, positive obligation, and remedies require reinterpretation in light of technological mediation. While the normative content of human rights remains intact, their implementation must reckon with emergent forms of opacity, diffusion, and anticipatory control. The systemic risk posed by AI is therefore not merely a risk to individual rights but a risk to the effectiveness and coherence of the entire human rights architecture.[26]

By articulating these implications, the article situates the debate on AI and human rights at a structural level. It moves beyond sectoral analyses and case-specific risk assessments to highlight how technological systems shape the conditions under which legal norms operate. The practical consequence is that policy prescriptions—whether transparency mandates, auditing regimes, or algorithmic impact assessments—must be informed by a nuanced understanding of systemic effects. Only through such an approach can human rights law preserve its normative and operational integrity in an era increasingly defined by probabilistic and automated governance.

## 5. Historical and Conceptual Context/Systemic Risks

To fully appreciate the systemic implications of artificial intelligence, it is instructive to situate the current transformation within the historical evolution of responsibility and justification in international human rights law. From the post-World War II codification of rights in the Universal Declaration of Human Rights (UDHR)[27] and the International Covenant on Civil and Political Rights, human rights frameworks have consistently relied on identifiable agents whose actions could be justified and reviewed. Responsibility was tied to human decision-making, and legal structures presupposed that harms could be traced, attributed, and remedied.[28] Positive obligations developed alongside this framework, requiring states to take proactive steps to prevent violations and to ensure effective remedies for affected individuals.[29]

Over the decades, the interpretation of responsibility has expanded in scope but remained grounded in human agency. For instance, supervisory or institutional accountability has been developed to address harms committed by state actors or delegated authorities, without displacing the principle that a responsible agent must ultimately justify their decisions. Similarly, remedies have evolved from basic reparations to more complex procedural guarantees, such as access to courts and independent oversight bodies, yet all remain premised on the ability to trace causality and exercise contestation. This conceptual lineage explains why AI introduces structural tension: the foundations of accountability, justification, and enforceability are predicated on assumptions that AI progressively challenges.

Artificial intelligence generates systemic risks that transcend individual sectors or rights.[30] Its probabilistic and anticipatory logic does not only affect discrete entitlements, but reshapes the broader architecture through which rights are

conceptualized and enforced. These systemic effects are particularly evident in domains such as governance, law enforcement, and social infrastructure.

In governance, AI-mediated decision-support systems influence policy-making by predicting social outcomes, optimizing resource allocation, or flagging potential risks. While enhancing efficiency, these systems can redistribute power in subtle ways, privileging those who control the algorithms or datasets. Normative authority becomes intertwined with technological design, subtly shifting accountability from human policymakers to socio-technical infrastructures.[31]

In law enforcement, predictive policing and risk assessment tools exemplify how AI alters the exercise of discretion. Decisions are filtered through probabilistic outputs, which can reproduce historical biases even when formal compliance frameworks are observed. Here, systemic risk manifests in the erosion of procedural fairness and the dilution of state accountability. The impact is not limited to individual cases, but extends to patterns of surveillance, enforcement priorities, and community trust, demonstrating a transversal effect across rights such as equality, privacy, and due process.[32]

Social infrastructure, including welfare systems, healthcare triage, and social services, is similarly affected. AI systems can determine eligibility, prioritize cases, or allocate scarce resources based on probabilistic scoring, producing outcomes that are difficult to contest. Systemic risks emerge not from isolated errors but from the cumulative and emergent properties of AI-mediated governance: opacity, adaptivity, and distributed decision-making challenge the traditional guarantees of human rights, including accessibility, fairness, and accountability.

This transversal perspective emphasizes that AI's most significant risks are systemic, rather than sectoral. By transforming the conditions under which decisions are made, responsibility is diffused, positive obligations are more complex to fulfill, and remedies are increasingly constrained. Legal scholars and policymakers must recognize that AI does not simply create new risks to existing rights; it reshapes the operational logic of the rights framework itself, requiring a reconceptualization of foundational principles.[33]

The extended theoretical and conceptual analysis demonstrates that artificial intelligence represents a structural challenge to the architecture of human rights law. By shifting decision-making from justification to probability, AI disrupts traditional assumptions about human agency, accountability, and remedial processes. This disruption is neither abstract nor hypothetical; it manifests in concrete applications across law enforcement, social media governance, and welfare administration.

Understanding AI as a systemic actor allows scholars to move beyond fragmented, sectoral analyses, offering a framework to assess the cumulative and structural effects of technological mediation. It also clarifies why conventional doctrinal remedies, compliance measures, and ex post adjudication are insufficient on their own. The theoretical lens adopted here provides both a diagnostic and a

conceptual tool, highlighting how existing norms may need reinterpretation to maintain coherence, enforceability, and legitimacy in an AI-mediated context.

These insights naturally lead to a synthesis of the central claim and to reflections on future research directions. The subsequent conclusion will integrate the arguments developed across introduction, methodology, state of the art, theoretical core, and systemic implications, and will outline the broader significance of reconceptualizing human rights in light of AI's structural influence.

**Conclusion**

This article has argued that artificial intelligence represents not merely a set of new tools or a collection of discrete risks for human rights, but a systemic transformation of the very architecture upon which rights are conceptualized, enforced, and defended. By shifting decision-making from justification to probability, AI challenges foundational assumptions regarding responsibility, positive obligations, and effective remedies. Traditional human rights frameworks, grounded in human agency and traceable causality, encounter tension when confronted with probabilistic, opaque, and adaptive systems.

The analysis demonstrates that these challenges are not limited to particular sectors or technologies. Rather, they are systemic, transversal, and cumulative. Examples drawn from law enforcement, social media governance, and welfare systems illustrate that AI-mediated decision-making diffuses responsibility, constrains remedies, and redefines the conditions under which rights can be effectively protected. The structural transformation affects both public and private actors, highlighting the inadequacy of sectoral or case-based approaches in capturing the full scope of AI's impact.[34]

This systemic perspective has important theoretical implications. Responsibility may need to be reconceptualized as distributed or layered, positive obligations as anticipatory and structural, and remedies as proactive and transparency-oriented. Such reconceptualizations do not abandon human rights law but demand a reinterpretation of its core principles in light of technological mediation. By foregrounding the shift from justification to probability, this article situates AI as a transformative force that reshapes the normative and operational conditions for human rights, rather than simply as a source of violations to be mitigated.

The originality of this contribution lies in its combination of critical theory, historical perspective, and practical illustration. By integrating doctrinal analysis with systemic insights, the article highlights emergent tensions that may otherwise remain obscured in sectoral studies. In doing so, it provides both a conceptual framework for understanding AI's structural effects and a diagnostic lens for assessing the coherence and enforceability of human rights norms under conditions of technological mediation.

Finally, the article opens avenues for further research. Scholars might explore how distributed responsibility frameworks could be operationalized, how systemic auditing could be reconciled with normative guarantees, or how AI reshapes other

foundational legal concepts, such as proportionality, equality, and procedural fairness. Importantly, future work should continue to bridge doctrinal analysis, empirical observation, and critical theory, ensuring that human rights scholarship remains attuned to the structural transformations introduced by emerging technologies. Only by understanding AI as a systemic actor, rather than a collection of isolated risks, can human rights law anticipate, interpret, and adapt to the challenges of an increasingly probabilistic world.

## Endnotes

[1] United Nations Office of the High Commissioner for Human Rights (OHCHR), *The Right to Privacy in the Digital Age*, Report of the Office of the United Nations High Commissioner for Human Rights, UN Doc. A/HRC/60/45 (Geneva: United Nations, 29 August 2025), https://undocs.org/A/HRC/60/45.

[2] Antoinette Rouvroy and Thomas Berns, "Algorithmic Governmentality and Prospects of Emancipation. Disparateness as a precondition for individuation through relationships?", English trans. by Liz Carey-Libbrecht, *Réseaux* 177 (2013): 163–196, https://doi.org/10.3917/res.177.0163.

[3] See, inter alia, Sandra Wachter, Brent Mittelstadt, and Luciano Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation," *International Data Privacy Law* 7, no. 2 (2017): 76–99, https://doi.org/10.1093/idpl/ipx005; Mireille Hildebrandt, *Law for Computer Scientists and Other Folk* (Oxford: Oxford University Press, 2020); Karen Yeung, "Algorithmic regulation: A critical interrogation", *Regulation & Governance* 12, no. 4 (2018): 505–523, https://doi.org/10.1111/rego.12158.

[4] Catelijne Muller, *The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law*, report prepared for the Ad hoc Committee on Artificial Intelligence (CAHAI), CAHAI(2020)06-fin (Strasbourg: Council of Europe, 2020), 12–15, paras. 43–57, https://rm.coe.int/cahai-2020-06-fin-c-muller-the-impact-of-ai-on-human-rights-democracy-/16809ed6da.

[5] UN General Assembly, Resolution 2200A (XXI), *International Covenant on Civil and Political Rights*, on 16 December 1966, entered into force 23 March 1976.

[6] Council of Europe, "Convention for the Protection of Human Rights and Fundamental Freedoms", opened for signature 4 November 1950, 213 *U.N.T.S.* 221; *European Treaty Series* no. 5.

[7] *Golder v. the United Kingdom*, no. 4451/70, 21 February 1975, European Court of Human Rights.

[8] Hildebrandt, *Law for Computer Scientists and Other Folk*.

[9] Yeung, *Algorithmic Regulation*.

[10] Roger Brownsword, *Law, Technology and Society: Reimagining the Regulatory Environment* (London: Routledge, 2019), https://doi.org/10.4324/9781351128186.

[11] Marco Marsili, "Guerre à la Carte: Cyber, Information, Cognitive Warfare and the Metaverse", *Applied Cybersecurity & Internet Governance* 2, no. 1 (2023): 105–120, https://doi.org/10.60097/ACIG/162861.

[12] Luciano Floridi et al., "AI4People—An Ethical Framework for a Good AI Society. Opportunities, Risks, Principles, and Recommendations", *Minds and Machines* 28, no. 4 (2018): 689–707, https://doi.org/10.1007/s11023-018-9482-5.

[13] Muller, *The Impact of Artificial Intelligence*.

[14] Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact", *California Law Review* 104, no. 3 (2016): 671–732, https://doi.org/10.15779/Z38BG31.

[15] Alessandro Mantelero, "AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment", *Computer Law & Security Review* 34, no. 4 (2018): 754–72, https://doi.org/10.1016/j.clsr.2018.05.020.

[16] See, inter alia, Luciano Floridi et al., AI4People; Sandra Wachter, Brent Mittelstadt, and Chris Russell, "Counterfactual Explanations without Opening the Black Box. Automated Decisions and the GDPR", *Harvard Journal of Law & Technology* 31, no. 2 (2018): 841–887.

[17] Yeung, *Algorithmic Regulation*.

[18] Louise Amoore, *The Politics of Possibility: Risk and Security Beyond Probability* (Durham: Duke University Press, 2013).

[19] Brownsword, *Law, Technology and Society*.

[20] Hildebrandt, *Law for Computer Scientists and Other Folk*.

[21] Hildebrandt, *Law for Computer Scientists and Other Folk*; Marco Marsili, "Military Emerging Disruptive Technologies: Compliance with International Law and Ethical Standards", in *Intelligent and Autonomous: Transforming Values in the Face of Technology*, eds. Ignas Kalpokas and Julija Kalpokienė (Leiden: Brill, 2023), 25-50, https://doi.org/10.1163/9789004547261_004.

[22] OHCHR, A/HRC/60/45; Muller, *The Impact of Artificial Intelligence*.

[23] Hildebrandt, *Law for Computer Scientists and Other Folk*.

[24] David Lyon, *The Culture of Surveillance: Watching as a Way of Life* (Cambridge: Polity Press, 2018).

[25] Mireille Hildebrandt, "Primitives of Legal Protection in the Era of Data-Driven Platforms", *Georgetown Law Technology Review* 2, no. 2 (2018): 252–273.

[26] Brownsword, *Law, Technology and Society*.

[27] UN General Assembly, Resolution 217A (III), *Universal Declaration of Human Rights*, A/RES/217(III) A, 10 December 1948.

[28] François Ewald, "Insurance and Risk", in *The Foucault Effect: Studies in Governmentality*, eds. Graham Burchell, Colin Gordon and Peter Miller (Chicago: University of Chicago Press, 1991), 197–210.

[29] Walter Kälin and Jörg Künzli, *The Law of International Human Rights Protection* (Oxford: Oxford University Press, 2009); Dinah Shelton, *Remedies in International Human Rights Law*, 3rd ed. (Oxford: Oxford University Press, 2015).

[30] Ulrich Beck, *Risk Society: Towards a New Modernity* (London: Sage, 1992).

[31] Julie E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford: Oxford University Press, 2019).

[32] Barocas and Selbst, *Big Data's Disparate Impact*.

[33] Yeung, *Algorithmic Regulation*; Brownsword, *Law, Technology and Society*.

[34] Shoshana Zuboff, *The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power* (New York: PublicAffairs, 2019).