



Received: 2025/09/03
Accepted: 2026/04/15
Published: 2026/06/20

1. Department of Law, Lenj.C.,
Islamic Azad University,
Lenjan, Iran. *. Corresponding
Author. Email:

mahsakhajeh1372@gmail.com

2. Department of Law, Na.C.,
Islamic Azad University,
Najafabad, Iran. Email:

masoudraei@yahoo.com

3. Department of Law, Ka.C.,
Islamic Azad University,
Karaj, Iran. Email:

mabdulmaleki@gmail.com

How to cite this article:
Khajeh, M., Raei, M., &
Abdolmaleki, M. (2026). AI
and Fourth Generation of
Human Rights Interaction. *The
International Journal of
Humanities* 33(2): 1–14.

RESEARCH ARTICLE

AI and Fourth Generation of Human Rights Interaction

Mahsa Khajeh^{1*} , Masoud Raei² , Mahdi Abdolmaleki³

Abstract

Artificial intelligence is evolving at a remarkable and accelerating pace. This technology is present or becoming present in all areas of human life. In the future, the relationship between artificial intelligence and humans, and how it will interact with human rights, has become a subject of serious scholarly and policy debate. The international community's solution to this potential challenge is to create new rights under the title of the "Fourth Generation of Human Rights." Artificial intelligence will become an intelligent actor like humans. Therefore, the fundamental question is whether there will be a constructive interaction between artificial intelligence and the fourth generation of human rights. To answer this question, this research employs a descriptive-analytical methodology, examining primary sources such as international legal instruments and seminal academic works to construct its argument. The research suggests that, although artificial intelligence has raised significant issues in this field, it should be noted that preserving human dignity and protecting humans against the risks of emerging technologies, as the foundation of human rights, is essential for ensuring a balanced coexistence between artificial intelligence and the principles of the fourth generation of human rights. This paper contributes to the literature by synthesizing the technological and epistemic strands of Fourth Generation theory and proposing a coherent framework for AI-human rights interaction based on core principles of epistemic agency, accountability, and human-centric governance. The term "Fourth Generation of Human Rights," although not yet the subject of a formal international consensus, is increasingly used in academic and policy discourse to conceptually frame the emerging rights and governance mechanisms required to address the socio-legal challenges of the digital age.

Keywords: Artificial Intelligence, Fourth Generation, Inherent Human Dignity, International Law, Human Rights

Introduction

Human rights, by the international community, were transformed into rules within documents, and in an academic and general classification, were categorized into three generations: first (civil and political rights), second (economic, social, and cultural rights), and third (solidarity

rights). Human life on Earth and its way of living is so dynamic that it is changing and transforming annually, monthly, daily, and now hourly. These changes, which currently are profoundly influenced by technology, have led to the emergence of many new phenomena such as “Metaverse” and “Artificial Intelligence.” In the past, humanity’s response to these needs was reflected in the three generations, commensurate with each issue. Now, according to many thinkers, the time has come that, considering the rapid growth of technology and its effects on human life, new rights should be created and declared in the form of a new generation to achieve two things.

First, since the core of human rights is the inherent dignity¹ of human beings, and all these rights, which have been transformed from rights to rules, are ultimately to ensure this dignity, it is necessary that this foundation and core be continuously protected. Second, human life is constantly changing, and this change requires the creation of new rights (if not the declaration of a right) to cover all emerging aspects. Simultaneously with these developments and evolutions, the international community has not forgotten humanity in the face of new technologies and has not left them alone amidst the multitude of technologies, including artificial intelligence, but has declared rights that are now referred to as the “Fourth Generation of Human Rights.”

Meanwhile, the issue of artificial intelligence, which is now one of the most important (if not the most important) technological issues, is the subject of many fears and hopes about its future and its role in human life. Therefore, the issue of the encounter between the fourth generation of human rights and artificial intelligence can be very worthy of consideration and reflection. This paper therefore addresses the central research problem of how a coherent conceptualization of the Fourth Generation can mitigate the novel human rights risks posed by AI’s unique characteristics, such as autonomous decision-making and predictive profiling. In this regard, this paper intends to examine the place and origin of artificial intelligence in relation to the rights of the fourth generation, while reviewing the conceptual evolution of previous generations of human rights. It also emphasizes the challenges and the necessity of creating new legal mechanisms for a responsible encounter with technological developments, including artificial intelligence.

The Emerging of Generations of Human Rights

The Universal Declaration of Human Rights in 1948 marked the beginning of the transformation of rights into rules. The declaration, without any specific categorization under any title such as generation or the like, codified human rights. However, it was after that declaration that, due to political and legal considerations arising from the Cold War, the rights enshrined in the Universal Declaration were separated and approved by governments in two separate binding documents, and human rights from then on became human rights rules. Accordingly, the first category of human rights, which are referred to by titles such as negative

¹ Human dignity relates to the recognition of the intrinsic and equal worth of each individual human being, regardless of race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other grounds (UNESCO, 2021, Para 13).

rights or individual rights against the state, came together in the International Covenant on Civil and Political Rights, and another category was included under the title of positive rights or state obligations in the International Covenant on Economic, Social and Cultural Rights.

Despite this, the concept of “generation” was not yet used to categorize human rights. Considering the emergence of new issues in the international arena as required by the times, the international community deemed other rights necessary for human “life,” which was generally either absent or less prominent in previous rights. These rights, which were collective in nature, relate to issues such as peace, environment, and sustainable development under the umbrella of solidarity rights. With the completion of the three sides of this triangle of human rights, the Czech-born jurist, Karel Vasak, inspired by the three principles of freedom, equality, and fraternity enshrined in the French Declaration of Human Rights (Risse, 2021, p. 11), first used the categorization of “generations of human rights” in the form of the first, second, and third generations. The rights of each of these generations were established as required by their time to meet the needs and enjoyment of a better life for humanity.

This proposition, namely the establishment of rights for the enjoyment of a better life, is felt more than ever in this age of technology, and is indeed a necessity. For this reason, some thinkers believe that the existing three generations of human rights are not adequate to meet the new needs commensurate with the new technologies. The novelty lies in technologies that exhibit autonomy (self-directed decision-making), self-evolution (learning and adapting beyond initial programming), and predictive capacity based on mass data analysis. These characteristics challenge core human rights principles like autonomy, non-discrimination, and due process in fundamentally new ways that existing frameworks struggle to address (Yeung, Howes, & Pogrebna, 2020). These generations not only surround human life but also have a direct impact on it, so another way must be sought for, and a new plan must be devised in the form of establishing new rights. Accordingly, recently, and not so long ago, the issue of a new generation of human rights, known as the “Fourth Generation,” has been raised, a generation that is, in a way, for “rethinking the human condition in a technological society” (Bustamante, 2001, p. 2) and focuses on the values of “identity” and “precaution” as reference values (Baroni, 2024, p. 39).

Fourth Generation Of Human Rights In The Context Of Technological Change

The existence or non-existence of a fourth generation of human rights is now a question. Whether such rights have fundamentally created and exist now, or whether they are in the process of created and emerging? Although some believe that there have been scattered discussions of the fourth generation since the 1980s (Risse, 2021, p. 8), it can be said that there is still no consensus on the formation of this new generation, and its content is not yet very clear. The use of the term “fourth generation” is thus primarily a heuristic and conceptual tool to organize discourse around novel, technology-related rights claims, rather than denoting a codified, universally accepted legal category. The term “fourth generation of human rights” is employed, even in the absence of formal international consensus, primarily as a conceptual and framing tool in academic and policy discourse. Its use offers several key benefits: First, it organizes the debate around novel rights claims and legal challenges arising from the digital

age. Second, it focuses attention from lawmakers, scholars, and international bodies on the urgent need to develop new legal and governance frameworks for emerging technologies like AI. Third, by establishing historical continuity with the previous three generations, it highlights the dynamic and evolving nature of human rights in response to profound societal transformations (Lubin, 2021; Baroni, 2024, p. 34; Risse, 2021). On the other hand, however, some consider the fourth generation to be formed or in the process of formation. They also consider its content to be clear, they consider the general content of this generation to include rights centered on technology, and they believe that the subject of the fourth generation can be the mixing of humans in the rich possibilities of the digital life world, a world that includes entities beyond human intelligence (Risse, 2021, p. 9).

In this sense, this generation implies rights that relate to the digital realm, and its scope is new technologies (Ma, 2023). In the meantime, however, there is an extraordinarily exciting theory that believes that the rights of the fourth generation are of the nature of epistemic and cognitive rights, and fundamentally, human rights as human rights, as they are applied in digital lives, should be considered as the next generation (Risse, 2021, p. 8). While emerging technologies may appear to challenge existing rights primarily at the level of implementation, this article argues that certain technological characteristics introduce qualitatively novel human rights concerns. In particular, opaque and autonomous algorithmic decision-making systems fundamentally alter the conditions under which individual autonomy is exercised. Unlike traditional forms of rights infringement, these systems can render the grounds of decision-making structurally inaccessible to affected individuals, thereby undermining their capacity to understand, contest, or meaningfully respond to decisions that shape their lives. Existing generations of human rights, which largely presuppose human decision-makers and identifiable loci of responsibility, are not fully equipped to address this form of epistemic displacement. Consequently, the challenges posed by self-learning and predictive AI systems necessitate a reconceptualization of human rights that integrates both technological specificity and epistemic protection, as articulated within the framework of a Fourth Generation of Human Rights (Bustamante, 2001; Risse, 2021; Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016). This view is important because it establishes a connection between human, knowledge, and the digital world. Therefore, “this requires the establishment of a new”digital human rights” concept, a corresponding mechanism for human rights protection, and the “moral infrastructure” that respects the value of human rights, so as to provide an effective legal protection for the “digital human rights”” (Yan & Jiang, 2024; Ma, 2023, p. 239).

After the first generation, which related to the protection of the person, the second generation, which related to human dignity, and the third generation, which related to various collective efforts, the fourth generation now relates to humanity (Risse, 2021, p. 9). In this regard, it can be said that this generation connects and is the intersection of ethics, humanity (human rights), and technology. While a single, universally accepted definition remains elusive, a comprehensive working definition for this emerging generation can be proposed as: a set of rights, principles, and governance frameworks designed to protect and promote human dignity, autonomy, and equality in the context of rapid digitalization and intelligent systems, with a

focus on epistemic agency, data sovereignty, algorithmic accountability, and the preservation of human-centric values in socio-technical environments (Baroni, 2024, p. 34; Bustamante, 2001; Risse, 2021; Yeung, Howes, & Pogrebna, 2020; Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016; Floridi, 2014). A critical distinction must be made between ethical principles and binding legal norms. While principles like transparency and fairness are foundational, the Fourth Generation project involves their translation into enforceable human rights obligations for states and non-state actors, moving from soft ethics to hard law (Yeung, Howes, & Pogrebna, 2020). Human rights, on the one hand, are ethical rights, not in the sense that they do not entail a legal obligation, but in the sense that, in principle, these rights are ethics-oriented. On the other hand, these rights have emerged to provide a life based on inherent dignity for human beings. Ethics and humanity, which are twins, are now living in an era, which is the age of technologies that can both threaten ethics and humanity and guarantee the promotion and improvement of ethics and the quality of human life.

New technologies are not of the type of technologies of the past decades but are of the kind of achievements that can themselves be creators, innovators, and decision-makers. With these interpretations, thinkers have been seriously calling for the necessity of creating rights appropriate to this era and its amazing technologies for more than a decade. This need has led the international community towards rights to regulate human life in accordance with the technological age. For this reason, there is talk of the fourth generation of human rights, which deals with the rights of the digital age and the emergence of super-technologies (Ekram, 2025, p. 75); because it seems that the current three generations of human rights cannot solve the challenges arising from technology, so it is necessary to create the fourth generation of human rights (Baroni, 2024, p. 39).

The Discourse of Artificial Intelligence and the Fourth Generation of Human Rights

The rights enshrined in the Universal Declaration of Human Rights are a collection of all individual, social, political, and economic rights, and there is no distinction between them in this regard. However, in other documents, categorizations related to the first and second generations of human rights, generally of the nature of individual rights, were created, namely the commitment of the state to individual human beings. Following these two generations, the third generation took shape with group and collective concepts related to collective human life, such as the environment, peace, and sustainable development. Now, the expansion of technology and the creation of new technologies have reached a point where it can undoubtedly be called the “Age of Technology.” Technologies like artificial intelligence, blockchain, the Internet of Things, and the metaverse are not only increasingly visible in human life but also are shaping it in a way. Nevertheless, we are currently in the midst of transitioning from a phase of the technological age that is considered the gateway to the digital age and the age of artificial intelligence. This introduction indicates that “the main components of the fourth generation should be the next steps for the development of the human rights project” (Risse, 2021, p. 9). Consequently, a generation of rights appropriate to this era must be formed, capable of fulfilling two main functions:

First, it should provide the ground for humanity to benefit from the technologies of this era to ensure a better life.

Second, it should protect the inherent dignity and rights of human beings against new technologies.

With reference to the two functions above, the fourth generation of human rights, based on human dignity in the digital age, as mentioned, must include epistemic and cognitive rights so that a balance can be established between artificial intelligence or the age of artificial intelligence and human life. For this reason, Mathias Risse suggests that “The upcoming list of rights should be understood cumulatively, in the sense that rights introduced to protect epistemic actors in one role might well also protect them in others, but I will not list them again. The most important addition to the list of epistemic rights that the human rights framework already contains is rights to protect individuals in their roles as parts of the collective epistemic object” (Risse, 2021, p. 14). Therefore, he introduces four types of rights as rights to protect involved individuals:

1. Rights to protect individuals as individual epistemic subjects (knowers) (Risse, 2021, p. 14).
2. Rights to protect individuals in their roles as belonging to the collective epistemic subject (Risse, 2021, p. 14).
3. Rights to protect individuals as individual epistemic objects (knowns) (Risse, 2021, p. 15).
4. Rights to protect individuals in their roles as belonging to the collective epistemic object (Risse, 2021, p. 15).

Artificial Intelligence: Non-Biological Cognition and the Origin of the Third Intelligence

Intelligence, according to what most psychologists agree upon, is generally the capacity for abstraction, logic, understanding, and learning, the ultimate result of which is the power of reasoning, inference, and deduction from data. Unlike some other abilities, such as the power of speech, intelligence is not exclusive to humans and is shared between animals and humans. Based on this, animals are also capable of learning and using what they have learned. Today, however, the world faces a third type of intelligence that, although human-made, relies on a third actor called the “machine,” which is also, coincidentally, a creation of humans. Human intelligence has created a new and emerging intelligence called “artificial intelligence,” which has the ability to learn, reason, and act outside the human brain and head, in computer memory. Artificial intelligence has been defined in various ways, from simple to technical definitions. For example, some consider it a technology that allows computers to act intelligently. Others view artificial intelligence as a machine that will replace human labor to work more efficiently and faster for people. Another group offers a more technical definition: artificial intelligence is a system that has the ability to correctly analyze external data, learn from such data, and use that learning to achieve specific goals and tasks through flexible adaptation (Kaplan & Haenlein, 2019, pp. 15–25). This definition, drawing from Russell and Norvig’s canonical work, provides a sufficient and precise technical foundation for the subsequent analysis, avoiding redundant definitions (Russell & Norvig, 2009).

Given the above definitions, artificial intelligence, in a simple and yet more complete sense than other definitions, is characterized as computer programs designed to solve complex

problems by implementations of processes similar to human cognition (Cheng-Tek Tai, 2020, p. 339; Russell & Norvig, 2009). This definition presents a machine-based intelligence similar to human perception, which “includes reasoning, understanding, organizing, studying, communication, vision, and the ability to manipulate and move objects” (Singh, 2019, p. 556). This level of cognitive, reasoning, and inferential ability for machines, regardless of its benefits, seems potentially capable of creating challenges, first for humans and then for what is uniquely human, such as human rights. This is because the fundamental basis of artificial intelligence was to place a uniquely human characteristic, or in other words, the fundamental human trait, namely reason, on a computer. However, this issue is not without philosophical objections, as some believe that “this presents philosophical problems concerning the nature of the reality and the behaviour of artificial entities” (Singh, 2019, p. 556).

Now that the machine is to be in the position of a wise human, fundamental questions arise: What is the relationship between this new intelligence and the fourth generation of digital-centric human rights? Is it possible to establish understanding between them?

AI's Systemic Threats to Human Dignity: Mechanisms, Typologies, and Asymmetric Impacts

The threat AI poses to human dignity is not abstract but operates through specific, interlinked socio-technical mechanisms that directly corrode the foundations of human agency and autonomy (Lubin, 2021). These mechanisms manifest in ways that existing human rights frameworks struggle to remedy due to their scale, opacity, and embeddedness in private infrastructure. Algorithmic bias and discrimination arise when historical data patterns or flawed model design lead to systematically unfair outcomes for certain groups. For instance, predictive policing algorithms can reinforce over-policing in marginalized neighborhoods, while automated hiring tools may disadvantage women or ethnic minorities. This violates the right to equality and non-discrimination by automating and scaling prejudice, often under a veil of technical neutrality (Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016). Surveillance and behavioral profiling, powered by mass data collection and analysis, transform individuals into transparent, predictable data subjects. This constant observability, exemplified by social credit systems or workplace monitoring AI, chills free expression, undermines privacy, and compromises mental autonomy—the ability to think and act free from undue external influence or prediction. It reduces human beings to manageable data points, directly attacking the core of personal dignity (Zuboff, 2019; Wachter, 2018). Perhaps most insidiously, cognitive manipulation and micro-targeting exploit detailed psychological profiles to deliver tailored content designed to subliminally influence beliefs, emotions, and behaviors. This threatens the integrity of thought and opinion formation—a core component of personal autonomy—and undermines the very possibility of informed consent and democratic deliberation (Susser, Roessler, & Nissenbaum, 2019). Collectively, these practices erode human agency—the capacity for meaningful self-determination—which is a bedrock of human dignity. They represent not merely new methods for violating old rights, but novel challenges that reframe the relationship between the individual, technology, and power. Therefore, a Fourth Generation

response must directly address these mechanisms by safeguarding epistemic agency—the right to know, understand, and contest the systems that shape one’s life (Risse, 2021).

Typologies of AI Risks and Asymmetric Impacts on Human Dignity

A nuanced analysis requires distinguishing between AI capabilities, as each category presents distinct challenges to human dignity and thus demands tailored regulatory approaches. Narrow or Specialized AI (e.g., facial recognition, automated decision-making in social benefits) poses acute, identifiable risks of discrimination, arbitrary interference, and mass surveillance. Its threats are often direct and measurable against existing rights benchmarks. Generative AI (e.g., large language models, deepfakes) introduces systemic risks to the epistemic environment. It can erode trust by amplifying disinformation, manipulating cultural and political discourse, and blurring the line between reality and fabrication, thereby destabilizing the shared information commons essential for individual and collective autonomy (Weidinger et al., 2024; Bender, Gebru, McMillan-Major, & Shmitchell, 2021). Artificial General Intelligence (AGI)—still a prospective technology—raises existential questions for human dignity, centering on value alignment and control. The core challenge is ensuring that a superintelligent system’s goals remain compatible with human survival, well-being, and moral values, guarding against the ultimate erosion of human agency (Bostrom, 2014; Russell, 2019). This typology clarifies that the Fourth Generation must be agile, containing both immediate protections against narrow AI harms and forward-looking governance principles for emerging capabilities.

The dignitary harms of AI are not distributed equally but are exacerbated by entrenched global power asymmetries. The phenomenon of “data colonialism” describes the modern practice whereby data is systematically extracted from populations—often in the Global South—and appropriated for value creation and governance by powerful corporate and state actors in technologically dominant regions (Couldry & Mejias, 2019). This creates new forms of dependency and digital inequality. For example, AI systems trained on biased or unrepresentative global datasets may fail or cause harm when deployed in local contexts, while the benefits of innovation accrue elsewhere. Furthermore, surveillance technologies exported from authoritarian states can reinforce oppression abroad (Taylor & Broeders, 2015). This global asymmetry means that the human rights vulnerabilities created by AI are profoundly reshaped, demanding an internationalist and decolonial approach within the Fourth Generation. It necessitates rights to data sovereignty, equitable access to AI benefits, and global cooperation to prevent a regulatory race to the bottom, ensuring that the digital age does not entrench a new form of human rights hierarchy (Šehović, 2020).

The Era of Artificial Intelligence: The Necessity of Revision of Human Rights

Our time, technologically, is laying the foundation for a path that will be the cornerstone of human digital-centric life in the future. The advent of the digital age brings technologies for the betterment of human life, such as artificial intelligence, which has the potential to replace humans. Although artificial intelligence is a human creation, its future remains uncertain, and the limits of its capabilities are yet to be determined. Therefore, the creation of new rights commensurate with it becomes necessary. In this context, one of the functions of the fourth generation of human rights, and in a way its goal, can be the protection of human life in the

presence of new forms of intelligence (Risse, 2021, p. 6). It is thus evident that these technologies can also have adverse effects and consequences on human rights. Therefore, rights must be established to protect humans from potential dangers. With this description, “No human being or human community should be harmed or subordinated, whether physically, economically, socially, politically, culturally or mentally during any phase of the life cycle of AI systems. Throughout the life cycle of AI systems, the quality of life of human beings should be enhanced, while the definition of ‘quality of life’ should be left open to individuals or groups, as long as there is no violation or abuse of human rights and fundamental freedoms, or the dignity of humans in terms of this definition” (UNESCO, 2021, Para 14). In simpler terms, the fourth generation of human rights (Yan & Jiang, 2024), in its first step, similar to paragraph (b) of Article 15 of the International Covenant on Economic, Social and Cultural Rights, seeks to realize the three fundamental pillars of human rights—respect, protection, and fulfillment—concerning rights related to new technologies. In its second step, it aims to ensure that these technologies, in the first instance, do not violate, the core of human rights, which is inherent human dignity, the foundation of the global, indivisible, and interdependent system of human rights and fundamental freedoms (UNESCO, 2021, p. 18), and in the second instance, do not violate other fundamental human rights. This is based on a clear understanding that, on the one hand, “New technologies need to provide new means to advocate, defend and exercise human rights and not to infringe them” (UNESCO, 2021, p. 10), and on the other hand, rights should ensure that innovations do not undermine the fundamental principles of dignity, fairness, and autonomy (Nuray, 2024). It should be noted that the violation of human rights is not limited to the violation of the first and second generations of rights, i.e., individual rights, but also includes the third generation, such as environmental rights, the right to peace, and the right to sustainable development, as well as the fourth generation.

These factors have led to the proposal for the development of principles for the bioethics of artificial intelligence. This can serve as a maximum barrier to prevent the violation of fundamental human rights. Michael Cheng-Tek Tai proposes four principles of bioethics for artificial intelligence (Cheng-Tek Tai, 2020, pp. 342–343). His four principles are:

- The first principle is beneficence. According to this principle, the purpose and functions of artificial intelligence should benefit all of human life, society, and the world. Artificial intelligence that could be destructive to the biosphere should be prohibited because its goal is the benefit of human society and therefore it should be altruistic, not self-centered.
- The second principle is Value upholding. This principle implies the compatibility of artificial intelligence with social values, and it should be such that the universal values governing the natural world order are respected. Artificial intelligence cannot rise above social and ethical norms.
- The third principle is lucidity. According to this principle, artificial intelligence should be transparent without concealing any hidden agendas.
- The fourth principle is accountability. This principle holds the designers and developers of artificial intelligence responsible for the outcome and impact that this technology has

on the entire human society and the world. Therefore, they must be accountable for everything they produce and create.

From Principles to Binding Governance: The Architecture of Contemporary International Law

The ethical principles discussed, while foundational, require translation into binding legal obligations to form the operational core of a Fourth Generation of Human Rights. This translation from *soft ethics* to *hard law* is precisely what recent landmark international instruments undertake. They do not merely suggest guidelines; they construct a concrete legal architecture that defines duties, establishes accountability, and creates enforcement pathways, thereby giving substantive content to the Fourth Generation framework (Lubin, 2021).

The EU AI Act (2024) represents the world's first comprehensive horizontal regulation of AI. Its risk-based taxonomy is a direct legal embodiment of the precautionary principle central to Fourth Generation discourse. By categorically prohibiting AI practices deemed to pose an “unacceptable risk” to fundamental rights (e.g., social scoring, real-time remote biometric identification in public spaces), it draws a red line that existing human rights law could not delineate with such specificity. For “high-risk” systems (e.g., in employment, education, or law enforcement), it imposes a rigorous set of ex-ante obligations, including conformity assessments, data governance, transparency, and human oversight. These are not optional ethics; they are enforceable legal duties on providers and deployers, backed by significant administrative fines, thus creating a powerful model for a rights-protective regulatory state in the digital age (European Parliament and Council, 2024).

The Council of Europe Framework Convention on AI (2024), as the first binding international treaty on the subject, moves the locus of obligation firmly to the state level. It codifies the due diligence obligation of states parties to adopt legislative measures that ensure AI systems within their jurisdiction are designed, developed, and used in a manner compatible with human rights, democracy, and the rule of law. This treaty explicitly connects the governance of AI to the entire lifecycle of a system, from design to decommissioning, mandating a holistic approach that previous human rights instruments, focused on state action against individuals, could not anticipate. Its significance lies in creating a common international legal floor for AI governance, directly operationalizing Fourth Generation concerns into treaty law (Council of Europe, 2024).

The UN AI Advisory Body's Final Report “Governing AI for Humanity” (2024) provides the indispensable global governance and normative framework. It authoritatively anchors the entire project of AI governance in the foundational principles of the UN Charter and international human rights law, arguing that these bodies of law are fully applicable to the digital realm. The report's call for inclusive global governance addresses the risk of fragmentation and “data colonialism” by advocating for multilateralism. Most importantly, it outlines a pathway for strengthened accountability, suggesting the potential need for new international institutions or the adaptation of existing ones (like the Human Rights Council's

mechanisms) to monitor and review AI-related human rights impacts, thereby proposing the institutional infrastructure for a global Fourth Generation response (United Nations, 2024).

Collectively, these instruments operationalize the accountability pillar of the Fourth Generation. They move beyond assigning abstract responsibility to specifying concrete duty-bearers and mechanisms: they clarify that states' due diligence obligations under international human rights law extend to taking proactive steps to regulate, monitor, and hold accountable private actors developing and deploying AI. They affirm that corporate responsibility to respect human rights, as outlined in the UN Guiding Principles, requires businesses to conduct Human Rights Due Diligence (HRDD) specifically for their AI activities. They mandate or encourage concrete oversight tools, such as mandatory, ex-ante Human Rights Impact Assessments (HRIAs) for high-risk AI systems—a procedural right that empowers regulators and potentially affected communities. Furthermore, they envision independent oversight bodies (e.g., AI Ombudsperson models or strengthened national human rights institutions) with the mandate to investigate grievances, provide redress, and advise governments, thus creating essential domestic implementation channels for Fourth Generation rights (Yeung, Howes, & Pogrebna, 2020).

In essence, these frameworks are not just *about* the Fourth Generation; they are its emerging *acquis juridique*—the living proof that the conceptual discourse is crystallizing into actionable, binding international and regional law. They translate the epistemic right to understanding (through assessments and transparency) and the right to remedy (through oversight bodies) into legal reality, providing the tangible “governance frameworks” called for in the working definition of the Fourth Generation.

Bioethical Principles for AI and Human-Centric Governance

Although these principles are merely proposed principles for the bioethics of artificial intelligence, their essence forms the core and central point of the fourth generation of human rights. The first principle indicates the adaptation of new technologies to human life. Moving from the individual (human) to an intermediate system called society and then reaching the world as a whole, this principle follows a logic that the first and second generations followed with respect to the individual and the third generation with respect to society and the world. The second principle, in the first instance, expresses the compatibility of technology with human rights, and in the second instance, the governance of universal values, a significant part of which are human rights, over artificial intelligence. The third principle, however, seeks to clarify a situation where artificial intelligence does not become a tool against humans. This means that its content should be clear and transparent and should not have any hidden agendas, because since artificial intelligence has the power of learning, reasoning, and inference (Intelligence, 2010), this capacity might be used against humans in a way that has caused some concern. Finally, the fourth principle, considering the profound impact of artificial intelligence on human society, issues a warning to designers not to pit humans against artificial intelligence and to carefully consider the consequences of their work and be accountable for the responsibilities arising from it. Perhaps a kind of obligation, which can be both from natural and legal persons, can be inferred from the fourth principle.

The message of these principles is “to protect, promote and respect human rights and fundamental freedoms, human dignity and equality, including gender equality; to safeguard the interests of present and future generations; to preserve the environment, biodiversity and ecosystems; and to respect cultural diversity in all stages of the AI system life cycle” (UNESCO, 2021, p. 18). In principle, “these principles acknowledge that AI governance does not take place in a vacuum, that international law, especially international human rights law, applies in relation to AI” (United Nations, 2024, p. 8).

Conclusion

Fundamental developments in artificial intelligence technology make a review of human rights norms inevitable. While artificial intelligence will greatly assist human life and its existence, there are concerns that it may also pose threats to humans and, in particular, their rights. Based on this, it seems that the existing legal systems are not adequate to address the dimensions of its potential challenges. Since artificial intelligence, as one of the most prominent manifestations of technological progress, is no longer merely a technological tool, but an intelligent and influential actor in social, economic, and even ethical structures. This highlights the necessity of developing new rights in the form of the fourth generation of human rights; a generation whose core and center is special attention to human dignity and will include the right to control data, protection of privacy, and technological accountability.

The fourth generation of human rights, which is the result of the evolution of the previous three generations, is considered an appropriate solution to gradually respond to the emerging needs of humanity. Artificial intelligence may be accompanied by complexities and challenges such as independent decision-making, lack of transparency, and unpredictable consequences, therefore, the need to identify and develop new rights, with a focus on new risks and the requirements of human techno-life, is felt more than ever. These challenges do not necessarily prevent interaction and understanding between artificial intelligence and the principles of human rights. On the contrary, by designing appropriate legal frameworks, adopting a human-centered approach in technology development, and clarifying the ethical and legal responsibilities of developers, the possibility of interaction between artificial intelligence and the fundamental principles of the fourth generation of human rights can be created. Ultimately, the realization of interaction and discourse between artificial intelligence and the fourth generation of human rights is not only possible, but also a historical necessity that requires attention to the inherent dignity of human beings. It is possible with the identification of new rights appropriate to techno-life, and the creation of global mechanisms for monitoring, guiding, and responding to the risks of emerging technologies. Therefore, it can be said that the fourth generation of human rights is not in conflict with technology, but in interaction with it, and its ultimate goal is to guide technology towards human ends. On the other hand, in the shadow of institutionalizing technological rights, one can hope for a future in which technology remains at the service of humanity.

This analysis contributes to the literature by synthesizing the technological and epistemic conceptualizations of the Fourth Generation into a coherent framework centered on protecting epistemic agency and human dignity. It grounds this framework in contemporary international

law and identifies critical implementation gaps. To move from discourse to action, future efforts must prioritize: a) enacting risk-based regulatory frameworks informed by instruments like the EU AI Act; b) institutionalizing mandatory AI Human Rights Impact Assessments; c) establishing specialized global and national oversight mechanisms; and d) developing redress pathways for victims of algorithmic harm, particularly from the Global South. Empirical research on real-world cases—such as discriminatory predictive policing or automated welfare denial—is essential to test and refine these proposed safeguards.

References

- [1] Baroni, M. J. L. (2024). Fourth generation human rights in view of the fourth industrial revolution. *Philosophies*, 9(2), 34.
- [2] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).
- [3] Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- [4] Bustamante, J. (2001). Hacia la cuarta generación de Derechos Humanos: repensando la condición humana en la sociedad tecnológica. *Revista Iberoamericana de Ciencia, Tecnología, Sociedad e Innovación*.
- [5] Cheng-Tek Tai, M. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal*, 32(4), 339–343.
- [6] Council of Europe. (2024, September 5). *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. Strasbourg.
- [7] Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- [8] Ekram, Y. M. (2025). The impact of artificial intelligence on the international human rights system. *Acta Globalis Humanitatis et Linguarum*, 2(2), 75.
- [9] European Parliament and Council. (2024, July 12). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L, 1–69.
- [10] Floridi, L. (2014). *The Fourth Revolution: How the infosphere is reshaping human reality*. Oxford University Press.
- [11] Intelligence, A. (2010). *Fundamentals of genetic algorithms artificial intelligence* (Thesis GA 42).
- [12] Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25.
- [13] Lubin, A. (2021). The fourth generation of human rights in the digital age. *Israel Law Review*, 54(3), 375–400.
- [14] Ma, C. (2023). The “Fourth Generation of Human Rights” under the background of the smart society and its protection. *Chinese Yearbook of Human Rights*, 5, 239.
- [15] Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2).
- [16] Nuray, A. (2024, December 24). Toward the fourth generation – AI and the future of human rights law. *Universal Legal Architect*.
- [17] Risse, M. (2021). *The fourth generation of human rights: Epistemic rights in digital lifeworlds* (Faculty Research Working Paper Series No. RWP21-027). Harvard Kennedy School.
- [18] Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.
- [19] Russell, S. J., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Prentice Hall.
- [20] Šehović, A. (2020). Reimagining data sovereignty from the ground up: Towards a vision for decolonial AI. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 702–703).

-
- [21] Singh, A. (2019). The concept of artificial intelligence. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 6(3), 556.
- [22] Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2).
- [23] Taylor, L., & Broeders, D. (2015). In the name of development: Power, profit and the datafication of the global South. *Geoforum*, 64, 229–237.
- [24] UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. General Conference of the United Nations Educational, Scientific and Cultural Organization, 41st session, Paris.
- [25] United Nations, Advisory Body on Artificial Intelligence. (2024, September). *Governing AI for humanity: Final report*. United Nations.
- [26] Wachter, S. (2018). Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer Law & Security Review*, 34(3), 436–449.
- [27] Weidinger, L., et al. (2024). Societal impacts of large language models. In *The Ethics of Artificial Intelligence* (pp. 145–179). Oxford University Press.
- [28] Yan, Z., & Jiang, S. (2024, April 15). The concept of digital human rights and its reshaping of basic rights. *The Journal of Human Rights*.
- [29] Yeung, K., Howes, A., & Pogrebn, G. (2020). AI governance by human rights–centered design, deliberation, and oversight. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford Handbook of Ethics of AI* (pp. 75–102). Oxford University Press.
- [30] Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.