

Lucia Bakošová

Rethinking Human Rights Protection in the Age of Artificial Intelligence



RETHINKING HUMAN RIGHTS PROTECTION IN THE AGE OF ARTIFICIAL INTELLIGENCE

Lucia Bakošová

leges

Citation:

BAKOŠOVÁ, Lucia, 2026. *Rethinking Human Rights Protection in the Age of Artificial Intelligence*. Praha: Leges, 2026.

Reviewers:

doc. JUDr. Mgr. Alla Tymofeyeva, Ph.D.

Právnická fakulta, Univerzita Karlova v Praze

Dr. Maruša Tekavčič Veber, PhD.

Faculty of Law, University of Ljubljana

Prof. Angelo Viglianisi Ferraro

Mediterranea University of Reggio Calabria



This publication was prepared as part of the research project “*Technological Progress and International Human Rights Law*” (No. VVGS-2023-2947), implemented within the Internal Scientific Grant System of Pavol Jozef Šafárik University in Košice. *Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V05-00008.*

Publisher: Nakladatelství Leges, s. r. o., Lublaňská 4/61, Prague, Czech Republic

Year of issue: 2026

Print run: 30 copies

Number of pages: 201, AA 12,9

First edition

Cover design Michaela Zoubek Vydrová

Layout Gradis

Print Tiskárna PROTISK, s. r. o., České Budějovice

www.knihyleges.cz

© Lucia Bakošová, 2026

ISBN print 978-80-7502-874-7

ISBN e-book 978-80-7502-875-4

KATALOGIZACE V KNIZE - NÁRODNÍ KNIHOVNA ČR

Bakošová, Lucia, 1993-

Rethinking human rights protection in the age of artificial intelligence / Lucia Bakošová. --

First edition. -- Prague : Leges, 2026. -- 1 online zdroj

Obsahuje bibliografii a bibliografické odkazy

ISBN 978-80-7502-875-4 (online ; pdf)

* 341.231.14 * 004.8 * 34 * (048.8) * (0.034.2:08)

– ochrana lidských práv

– umělá inteligence -- právní aspekty

– monografie

– elektronické knihy

341 - Mezinárodní právo [16]

CONTENTS

Preface	11
Acknowledgments	13
Abbreviations	15
Introduction	17
I. International Human Rights Law in the Age of Artificial Intelligence	25
1. Artificial Intelligence: Conceptual and Legal Definitions	26
2. The Emergence and Nature of International Human Rights Law	30
3. Normative Framework of International Human Rights Law	36
3.1. Universal Normative Framework of Human Rights Protection	36
3.2. Regional Human Rights Framework	39
3.2.1. European Human Rights Framework	39
3.2.2. African Human Rights Framework	42
3.2.3. Inter-American Human Rights Framework	43
4. Institutional Framework for the Protection of Human Rights	45
4.1. The United Nations Human Rights Protection Mechanism	45
4.1.1. Charter-Based Mechanisms: Human Rights Council and the Office of the United Nations High Commissioner for Human Rights	46
4.1.2. Treaty-Based Mechanisms: UN Human Rights Treaty Bodies	50
4.2. Regional Institutional Mechanisms for the Protection of Human Rights	53
4.2.1. European Human Rights Protection Mechanisms	54
4.2.2. African Human Rights Protection Mechanisms	57
4.2.3. Inter-American Human Rights Protection Mechanisms	59

5. Contemporary Challenges and the Impact of Technological Change	61
5.1. Structural Limitations of International Human Rights Law	61
5.2. Opacity and Algorithmic Decision-Making	62
5.3. Accountability and Governance Gap	64
5.4. Institutional and Operational Challenges	65
5.5. Role of Private Actors	67
II. Regulatory Responses – From Soft Law to Binding Frameworks on Artificial Intelligence	69
6. Soft Law Approaches to the Regulation of Artificial Intelligence	71
6.1. The United Nations	72
6.1.1. Human Rights-Based Approaches to AI Governance within the UN System	76
6.1.2. The United Nations Educational, Scientific and Cultural Organization	80
6.1.3. The United Nations Children’s Fund	83
6.2. The Organization for Economic Co-operation and Development	89
6.3. The Regional Non-binding Approaches to Artificial Intelligence	95
6.3.1. The African Perspective on Artificial Intelligence Regulation	96
6.3.2. The Organization of American States	100
7. Regional Binding Approaches to Artificial Intelligence	103
7.1. The European Union’s Regulatory Framework for Artificial Intelligence	104
7.1.1. The Artificial Intelligence Act and the Protection of Human Rights	108
7.2. The Council of Europe’s Regulation on Artificial Intelligence	115
7.2.1. Normative Foundations of the Framework Convention: The Role of Soft Law	116
7.2.2. The Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law	120

III. Institutional Frameworks for the Protection of Human Rights in the Age of Artificial Intelligence	131
8. The United Nations Human Rights System and Artificial Intelligence	132
8.1. The Emergence of AI-Specific Mechanisms within the United Nations	132
9. Regional Human Rights Mechanisms	138
9.1. The European Union Institutional Framework for Artificial Intelligence and Fundamental Rights	138
9.1.1. Governance Architecture under the Artificial Intelligence Act	141
9.1.2. Complementary EU Governance Structures Supporting Fundamental Rights Protection	143
9.2. Human Rights Protection and Emerging AI Governance within the Council of Europe	147
9.2.1. Governance and Oversight under the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law	150
9.3. The African Institutional Framework for Human Rights and Artificial Intelligence	152
9.4. The Inter-American Institutional Framework for Human Rights and Artificial Intelligence	154
IV. Emerging Digital and AI-Related Rights: Beyond the Expansion of Existing Protection	157
10. Emerging Dimensions of Human Rights in the Digital Age ..	159
10.1. AI-Related Rights and Safeguards	162
10.1.1. Right Not to Be Subjected to Algorithmic Manipulation	165
10.1.2. The Right to a Human Decision and a Human-To-Human Interaction	167
V. Conclusions and Future Perspectives	173
Reference list	181
About the author	201

PREFACE

This monograph examines the interaction between technological progress and international human rights law, with particular attention devoted to the development and use of artificial intelligence and its implications for the future of human rights protection. The increasing integration of artificial intelligence into decision-making processes across a wide range of societal domains has significantly altered the conditions under which human rights are exercised and protected. These developments raise fundamental legal questions concerning responsibility, accountability, and the ability of existing international legal frameworks to respond to technologically mediated forms of power.

The book is grounded in the author's long-term research interest in the adaptability of international law to structural change. Previous research has focused on situations in which international legal frameworks are confronted with sudden or complex challenges requiring flexible interpretation and institutional coordination. Against this background, artificial intelligence is approached not merely as a new regulatory object, but as a phenomenon that tests the resilience and coherence of international human rights law as a normative system.

At the international level, responses to artificial intelligence have so far been characterised by fragmentation and a strong reliance on non-binding instruments. While numerous initiatives emphasise the importance of human rights, democracy, and the rule of law, their practical impact remains uncertain in the absence of clear obligations and enforcement mechanisms. This monograph therefore treats artificial intelligence as a case study through which broader questions concerning the role of soft law, the interpretation of existing human rights norms, and the evolution of international protection mechanisms can be examined.

The research presented in this book was conducted within the framework of the project "Technological Progress and International Human Rights Law". The central objective of this project is to assess whether the current system of international human rights law provides sufficient normative and institutional tools to address the challenges posed by artificial intelligence, or whether these challenges require new approaches at the international level. Emphasis is placed on the interaction between binding legal norms and emerging regulatory initiatives at both the universal and regional levels.

A further focus of the monograph lies in the ongoing debate on the emergence of new human rights in response to technological change. Rather than presuming a rupture with established human rights doctrine, the analysis adopts a cautious approach that prioritises the reinterpretation and strengthening of existing guarantees. In doing so, the book reflects the view that international human rights law has historically evolved through gradual normative development, shaped by practice, institutional dialogue, and changing social realities.

This monograph is addressed primarily to scholars and students of international law, human rights practitioners, and institutional actors engaged in the governance of emerging technologies. By situating artificial intelligence within the broader framework of international human rights law, it aims to contribute to current debates on the future direction of human rights protection in an increasingly digital and automated environment.

Lucia Bakošová

ACKNOWLEDGMENTS

Some parts of this monograph have previously been published in the form of scientific articles and contributions in peer-reviewed collections and journals, serving as a basis for scholarly discussion and feedback on interim research findings. These sections have since been substantially revised, updated, and integrated into a coherent monographic framework, considering recent developments in the international human rights law and the legal regulation of artificial intelligence. The author would like to express her sincere appreciation to the reviewers of this monograph for their valuable comments and constructive feedback.

The author also acknowledges the use of generative artificial intelligence tools for limited language editing and stylistic refinement of the manuscript. The research, analysis and conclusions presented in this work remain entirely the author's own.

Dedicated with love to Juliána

ABBREVIATIONS

ACHPR	African Commission on Human and Peoples' Rights
AI	Artificial Intelligence
Art.	Article
AU	African Union
CAHAI	Ad Hoc Committee on Artificial Intelligence
CAI	Committee on Artificial Intelligence
CDDH	Steering Committee for Human Rights
CDDH-IA	Drafting Group on Human Rights and Artificial Intelligence
CJEU	Court of Justice of the European Union
COBRA	Context-Based Risk Analysis
COE	Council of Europe
EAIB	European Artificial Intelligence Board
ECHR	European Convention on Human Rights and Fundamental Freedoms
ECtHR	European Court of Human Rights
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EU	European Union
FRA	Fundamental Rights Agency
FRIA	Fundamental Rights Impact Assessment
GDPR	General Data Protection Regulation
GPAI	General-Purpose AI
HRC	Human Rights Council
HUDERIA	Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law
ICCPR	International Covenant on Civil and Political Rights
ICESCR	International Covenant on Economic, Social and Cultural Rights

IHRL	International Human Rights Law
ILC	International Law Commission
ITU	International Telecommunication Union
NGOs	Non-Governmental Organisations
OAS	Organization of American States
OECD	Organization for Economic Cooperation and Development
OHCHR	Office of the United Nations High Commissioner for Human Rights
SMEs	Small and Medium Enterprises
STC-CICT	Specialized Technical Committee on Communication and Information and Communication Technology
UDHR	Universal Declaration of Human Rights
UN	The United Nations
UNESCO	The United Nations Educational, Scientific and Cultural Organization
UNICEF	The United Nations Children's Fund

INTRODUCTION

Technological progress has long posed structural challenges for public international law, driving the continuous adaptation of its normative frameworks. Historically, such challenges have acted as catalysts for legal development, prompting international law to redefine its scope, principles, and institutional responses considering technological change. The advent of nuclear weapons, for instance, prompted the development of complex non-proliferation regimes and challenged the adequacy of existing rules of international humanitarian law. Similarly, the emergence of aviation and space technologies necessitated the articulation of new legal regimes governing airspace sovereignty and the use of outer space. More recently, the rise of the internet and digital technologies has generated novel questions concerning jurisdiction, state responsibility, and the applicability of international law in cyberspace. In parallel, technological advances in industrial production and biotechnology have significantly influenced the evolution of international environmental law, particularly regarding transboundary harm, risk regulation, and precautionary approaches.

Against this broader trajectory of technological evolution, artificial intelligence (hereinafter “AI”) emerges as a particularly transformative development. Unlike earlier technologies, AI systems are characterised by their capacity for autonomous decision-making and their integration into core societal functions, thereby challenging the adequacy of existing normative frameworks of public international law. AI systems increasingly influence a wide range of societal domains, from economic activity and public administration to environmental governance. In this context, their growing deployment across both public and private sectors has significant implications for the protection of human rights and fundamental freedoms. On the one hand, AI systems have the capacity to enhance communication, thereby contributing to the exercise of the rights to freedom of expression and association. More broadly, digital and AI-driven technologies may empower individuals and groups by improving the effectiveness and accessibility of public services, strengthening democratic participation, supporting civil society engage-

ment, and enabling more effective reporting, monitoring, and prevention of human rights violations.¹

However, the deployment of AI and other digital technologies also gives rise to a broad spectrum of risks for the protection of human rights. These include threats to the right to privacy resulting from extensive data collection and processing, as well as concerns related to cybersecurity and the integrity of digital systems. The increasing reliance on algorithmic decision-making further raises questions regarding the quality and authenticity of information, including the spread of misinformation and manipulation of public discourse,² thereby raising broader questions about human dignity, individual autonomy, and the resilience of the information environment. In addition, AI systems may reinforce existing patterns of discrimination,³ contribute to social polarization and inequality, and exacerbate the digital divide. More intrusive forms of surveillance,⁴ including mass and workplace monitoring,⁵ pose challenges to fundamental freedoms, while the misuse of digital technologies has enabled new forms of abuse, including cyberviolence and online

¹ UN Human Rights Council: *Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/47/52, 19 May 2021, para. 12; Council of Europe Commissioner for Human Rights: *The Human Line: Safeguarding Rights and Democracy in the AI Era*, CommHR(2025)57, 20 October 2025. Online: <https://rm.COE.int/commhr-2025-57-report-the-human-line-safeguarding-rights-and-democracy/48802906bd>; UN Office of the High Commissioner for Human Rights: *High Commissioner Türk: Human rights are the foundations for AI that benefits humanity: Remarks to AI Impact Summit plenary*, 19 February 2026, Delhi. Online: <https://www.ohchr.org/en/statements-and-speeches/2026/02/high-commissioner-turk-human-rights-are-foundations-ai-benefits>.

² VAN GILS, F., MÜLLER, W., PRÜFER, J.: *Microtargeting, voters' unawareness, and democracy*. In: *The Journal of Law, Economics, and Organization*, Vol. 41, No. 2, 2025, pp. 634–653.

³ MIHAJLOVIĆ, A., ĆORIĆ, V.: *Artificial Intelligence and Discrimination: Strengths and Weaknesses of the Current European Anti-Discrimination Legal Framework*. In: *Regional Law Review*, Vol. 5, 2024, pp. 9–25.

⁴ MARTYNOVA, E.: *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary*. In: BELLI, L., GASPAR, W. (eds.): *AI from the Global Majority: Official Outcome of the UN IGF Data and Artificial Intelligence Governance Coalition*. Rio de Janeiro: FGV Direito Rio, 2024, pp. 109–120.

⁵ ALOISI, A., GRAMANO, E.: *Artificial Intelligence Is Watching You at Work. Digital Surveillance, Employee Monitoring, and Regulatory Issues in the EU Context*. In: *Special Issue of Comparative Labor Law & Policy Journal*, “Automation, Artificial Intelligence and Labour Protection”, Vol. 41, No. 1, 2019, pp. 95–121.

harassment,⁶ as well as risks to the right to a fair trial arising from automated decision-making without adequate safeguards.⁷ Importantly, these risks are no longer merely hypothetical, but have already materialised in practice, prompting increasing attention at both universal and regional levels.

Given the inherently cross-border nature of digital technologies and the global reach of AI systems, national legal frameworks alone often prove insufficient to address the full range of challenges associated with these technologies. These developments underline the growing need for international legal regulation capable of ensuring that the development and deployment of AI systems take place in accordance with established standards of international human rights law (hereinafter “IHRL”). At the same time, from the perspective of effective global governance, it is important to avoid excessive fragmentation of legal regimes across different states and regions and to promote a degree of coherence in emerging regulatory approaches.

In this context, the increasing need for coherent international regulation has placed international intergovernmental organisations at the forefront of efforts to shape the legal and ethical governance of AI. This role is reflected in the activities of organisations such as the United Nations (hereinafter “UN”), the United Nations Educational, Scientific and Cultural Organization (hereinafter “UNESCO”), the European Union (hereinafter “EU”), the Council of Europe (hereinafter “COE”), and the Organisation for Economic Co-operation and Development (hereinafter “OECD”), which have contributed to the emerging regulatory landscape through the development of universal and regional principles, as well as both binding and non-binding frameworks for the trustworthy and human rights compliant use of AI systems. These initiatives illustrate the gradual emergence of a multi-layered in-

⁶ UN Human Rights Council: *Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/47/52, 19 May 2021, paras. 15–28; BENKO, R.: *The Impact of Artificial Intelligence on Fundamental Rights and Freedoms and Vice-Versa*. In: KLUČKA, J., BAKOŠOVÁ, L. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Nakladatelství Leges, 2021, p. 42.

⁷ TEO, S.A.: *Artificial Intelligence and Transitional Justice: Framing the Connections*, Impunity Watch 2025, pp. 7–46. Online: <https://www.impunitywatch.org/wp-content/uploads/2025/01/Overview-Paper-Report-Artificial-Intelligence-and-Transitional-Justice-Framing-the-Connections-Impunity-Watch-Jan-2025.pdf>; BALCERZAK, M.: *Introduction: AI and International Human Rights Law*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J.: *Artificial Intelligence and International Human Rights Law*, Northampton: Edward Elgar Publishing, 2024, pp. 1–4.

ternational governance framework for AI. Among the most influential are the *OECD Principles on Artificial Intelligence*,⁸ the *UNESCO Recommendation on the Ethics of Artificial Intelligence*,⁹ the EU's *Artificial Intelligence Act*,¹⁰ and the *COE Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*.¹¹ At the same time, regional organisations such as the Organization of American States (hereinafter "OAS"), or the African Union (hereinafter "AU") are increasingly engaging with the governance of AI. While their approaches remain at varying stages of development, these organisations have begun to articulate strategic frameworks, policy guidelines, and ethical principles reflecting regional priorities and levels of technological capacity.¹² This trend highlights the gradual expansion of AI-related norm-setting beyond traditionally dominant actors and underscores the emerging diversity of regulatory approaches within the global landscape of AI governance. Against this background, a structured legal analysis is required to assess both the opportunities and risks associated with AI systems, as well as the capacity of existing international legal frameworks to respond to these developments.

Methodologically, the monograph combines legal interpretation, comparative analysis, and doctrinal assessment. It examines relevant international treaties, soft-law instruments, jurisprudence, and institutional practice at both universal and regional levels. Emphasis is placed on the role of international organisations, supervisory bodies, and emerging advisory mechanisms in shaping the normative environment surrounding AI. The analysis also explores the interplay between binding and non-binding instruments,

⁸ OECD: *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, 2025.

⁹ UNESCO: *Recommendation on the Ethics of Artificial Intelligence*, SHS/BIO/PI/2021/1, 23 November 2021.

¹⁰ *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)*, 12 July 2024, OJ L 2024/1689 (hereinafter "AI Act").

¹¹ *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, Strasbourg, 17 May 2024 (hereinafter "Framework Convention").

¹² AU: *Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity*, July 2024. Online: https://au.int/sites/default/files/documents/44004-doc-EN-Continental_AI_Strategy_July_2024.pdf; OAS: *Inter-American Guidelines on Data Governance and AI*, 2024. Online: <https://www.oas.org/ext/DesktopModules/MVC/OAS-DnnModules/Views/Item/Download.aspx?type=1&id=1080&lang=1>.

paying specific attention to the capacity of soft law to function as a laboratory for the progressive development of IHRL.

The main objective of this monograph is to assess whether the existing system of international human rights law provides an adequate normative and institutional framework for addressing the human rights implications of the development and use of artificial intelligence, and to identify the forms of legal and institutional adaptation required where existing protection proves insufficient. Accordingly, the monograph addresses two principal research questions. First, to what extent can existing international human rights norms, interpretative approaches and institutional mechanisms respond effectively to the distinctive risks associated with AI systems? Second, do these risks necessitate the development of new international legal rules, new human rights or AI-specific institutional mechanisms?

The central thesis advanced in this monograph is that international human rights law remains the indispensable normative foundation for the governance of AI and is capable of addressing a substantial part of AI-related risks through the dynamic interpretation and effective implementation of existing rights. At the same time, the specific characteristics of AI systems, including their opacity, autonomy, transnational operation and dependence on complex networks of public and private actors, expose significant procedural, institutional and accountability gaps. Addressing these gaps does not require the replacement of the existing human rights framework, but rather its adaptation through stronger procedural safeguards, enhanced institutional capacity and coordination, clearer allocation of responsibility, and the cautious recognition of new rights or mechanisms where existing guarantees cannot adequately protect distinct normative interests.

Addressing these questions requires moving beyond a purely descriptive account of technological developments and regulatory initiatives. The monograph therefore adopts a critical legal approach that situates AI within the broader evolution of IHRL and examines its capacity to adapt through interpretation, institutional practice and normative innovation. Rather than presuming that technological novelty automatically demands new rights or legal instruments, the analysis identifies the points at which existing guarantees remain effective, where they encounter interpretative or operational limits, and where genuine normative gaps may arise.

The book further engages with the increasingly prominent discourse on so-called “emerging” or “new” human rights in the context of digitalisation and the use of AI. Proposals such as the right not to be subject to automat-

ed decision-making,¹³ the right not to be measured, analysed, or guided,¹⁴ the right to meaningful human contact,¹⁵ or the right to (re)start with clean (digital) slate¹⁶ reflect growing concerns regarding the impact of algorithmic systems on human autonomy and dignity. The monograph examines the international legal status, conceptual coherence and relationship of these proposals to established human rights, with particular attention to whether they constitute genuinely new rights or rather the specification, reinterpretation or extension of existing guarantees in response to evolving technological conditions.

Structurally, the monograph is divided into five parts. Following this introduction, Part I lays the conceptual and normative foundations by defining AI and situating it within the framework of IHRL. It examines the nature, sources, and institutional structure of IHRL, before addressing the contemporary challenges posed by technological change, including both conceptual and operational gaps.

Part II explores the regulatory responses to AI, moving from universal soft law initiatives developed within the UN system to regional approaches and governance models. Attention is devoted to both non-binding frameworks adopted at the regional level, including those emerging within the African and Inter-American contexts, and binding regulatory models developed by the EU and the COE. The analysis further examines the gradual transition from ethical principles and voluntary standards towards increasingly institutionalised and legally binding approaches to AI governance, with specific emphasis on their implications for the protection of human rights, democracy, and the rule of law.

Part III focuses on the institutional dimension of human rights protection in the context of AI. It analyses the role of UN human rights bodies, including treaty bodies and special procedures, as well as regional mech-

¹³ ROIG, A.: *Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR)*. In: *European Journal of Law and Technology*, Vol. 8, No. 3, 2017, pp. 1–17.

¹⁴ BRITO, I.: *The right not to be monitored: in defence of one of the last bastions of the human being against the emerging omniscient surveillance society*. In: *UNIO – EU Law Journal*, Vol. 11, No. 1, June 2025, pp. 52–87.

¹⁵ SHANY, Y.: *A Right to a Human-to-Human Interaction*, University of Oxford – Institute for Ethics in AI, 7 November 2024. Online: <https://www.oxford-aiethics.ox.ac.uk/blog/right-human-human-interaction>.

¹⁶ CUSTERS, B.: *New digital rights: Imagining additional fundamental rights for the digital era*. In: *Computer Law & Security Review*, Vol. 44, 2022, 105636, pp. 9–10.

anisms, and considers the emergence of AI-specific governance and advisory structures. Attention is devoted to the interaction between traditional human rights mechanisms and newly emerging regulatory and supervisory frameworks addressing AI-related risks. This part also reflects on the institutional and operational challenges that existing mechanisms may face in responding effectively to the increasing complexity and transnational nature of AI systems.

Part IV addresses the emerging dimensions of human rights in the digital age. It considers how digitalisation and AI challenge traditional understandings of human rights and raise questions about the adequacy of existing forms of protection. Rather than treating digital and AI-related rights as a closed category, it examines them as evolving normative claims that reflect broader concerns about autonomy, identity, participation, accountability, safety and trust in technologically mediated environments.

The final part synthesises the main findings of the monograph and formulates proposals *de lege ferenda* aimed at strengthening the protection of human rights in an era of accelerating technological transformation, including reflections on the potential recognition and consolidation of new human rights standards in response to the challenges posed by AI.

By examining AI through the lens of IHRL, this monograph seeks to contribute to the ongoing scholarly and policy debate on the governance of emerging technologies. Its aim is not only to identify risks and challenges, but also to offer a structured legal assessment of continuity and change within IHRL and to explore pathways for its effective adaptation to the realities of the algorithmic age.

I. INTERNATIONAL HUMAN RIGHTS LAW IN THE AGE OF ARTIFICIAL INTELLIGENCE

IHRL represents one of the most developed normative frameworks of contemporary international law, designed to protect the dignity, autonomy, and fundamental freedoms of individuals against the exercise of public power. Emerging in the aftermath of the Second World War and gradually evolving through a dense network of international treaties, customary norms, and institutional mechanisms, IHRL has traditionally operated within a paradigm centred on state responsibility and territorially bounded forms of governance.

However, the accelerating pace of technological development – and particularly the rise of AI systems – challenges some of the foundational assumptions upon which this framework was constructed. AI systems introduce novel forms of decision-making that are data-driven, increasingly autonomous, and often opaque, while their deployment frequently transcends national borders and involves a complex interplay between public authorities and private actors. As a result, the application of existing human rights norms can no longer be taken for granted, but instead requires careful reinterpretation and, in some instances, conceptual rethinking.

Against this background, this part of the monograph seeks to examine the capacity of IHRL to respond to technological change. It first outlines the conceptual foundations, sources, and institutional structure of the international human rights system, before turning to the structural limitations that arise when this system is confronted with algorithmic forms of governance. Rather than offering a comprehensive analysis of specific regulatory instruments, which are addressed in subsequent parts of the monograph, the aim here is to establish a normative and analytical framework that will guide the examination of AI-related human rights challenges in the chapters that follow. At the same time, such an assessment necessarily presupposes a clear understanding of what constitutes AI in both conceptual and legal terms. For this reason, the analysis begins with a clarification of the notion of AI, its defining characteristics, and the challenges associated with its legal qualification, which form the subject of the following chapter.

1. ARTIFICIAL INTELLIGENCE: CONCEPTUAL AND LEGAL DEFINITIONS

The concept of AI was formally introduced at the 1956 Dartmouth Summer Research Project on Artificial Intelligence, which is widely regarded as a foundational moment in the development of the field.¹⁷ Despite this early conceptualisation, the development of AI did not follow a linear trajectory. Over the following decades, the field experienced alternating periods of heightened expectations and subsequent stagnation, often referred to as “AI winters”. Early approaches to AI, particularly in the 1950s and 1960s, were primarily based on symbolic reasoning and rule-based systems, which aimed to replicate human problem-solving through formal logic and pre-defined rules. While these systems demonstrated promising results in controlled and narrowly defined tasks, they proved limited in their ability to operate in complex, real-world environments. Similar limitations persisted in later phases, including the development of expert systems in the 1980s, which, despite their practical applications, remained rigid and dependent on manually encoded knowledge.¹⁸

More recent developments in AI have been closely associated with advances in deep learning, which have significantly enhanced the capacity of systems to process large volumes of data and to perform complex tasks with a high degree of accuracy. In particular, the success of supervised learning approaches has contributed to the widespread deployment of AI systems across various domains. At the same time, further progress is increasingly expected to arise from unsupervised and reinforcement learning techniques, which enable systems to identify patterns, adapt to new environments, and make decisions with limited human intervention. In areas such as computer vision and natural language processing, these developments have already led to substantial improvements in performance, with AI systems demonstrating the ability to analyse images, understand text, and interact with users

¹⁷ MCCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C.: *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, August 31, 1955. In: *AI Magazine*, Vol. 27, No. 4, 2006, pp. 12–13.

¹⁸ NILSSON, N. J.: *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press, 2010, pp. 19–657; HAENLEIN, M., KAPLAN, A.: *A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence*. In: *California Management Review*, Vol. 61, No. 4, 2019, pp. 2–6.

in increasingly sophisticated ways. More broadly, the combination of deep learning with reinforcement learning and other advanced techniques points towards a new generation of AI systems capable not only of recognising patterns, but also of engaging in more complex forms of reasoning and decision-making.¹⁹ In this context, particular attention must be paid to the rapid emergence of generative AI, which is typically built upon large-scale foundation models trained on vast datasets and capable of performing a wide range of tasks, including the generation of text, images, audio, and code.²⁰ Unlike earlier systems primarily designed for classification or prediction, generative AI introduces qualitatively different capabilities, enabling the creation of outputs that may closely resemble human expression and creativity. This evolution further accentuates the shift from predictable, rule-based systems towards adaptive and often opaque models, thereby intensifying existing concerns related to transparency, accountability, and control, which are central to the protection of human rights in the digital age.

As the foregoing overview suggests, the understanding of what constitutes AI has evolved significantly over time, reflecting both technological developments and shifting conceptual approaches. This evolution is also mirrored in the diversity of definitions of AI across natural, technical, and legal sciences, underscoring its inherently interdisciplinary character. As a field of inquiry, AI brings together insights from computer science, mathematics, logic, cognitive science, psychology, and philosophy, among others.²¹ At the same time, AI may be understood from two interconnected perspectives. First, it denotes a research domain aimed at understanding, modelling, and replicating forms of intelligence and cognitive processes through computational and related methods.²² In this sense, it has been described as the science of enabling machines to perform tasks that would require intelligence if carried out by humans. Second, AI also refers to the tangible outcomes of this research,

¹⁹ LECUN, Y., BENGIO, Y., HINTON, G.: *Deep Learning*. In: *Nature*, Vol. 521, 2015, p. 442.

²⁰ BOMMASANI, R. et al.: *On the Opportunities and Risks of Foundation Models*, Center for Research on Foundation Models, 2021, pp. 3–7. Online: <https://crfm.stanford.edu/assets/report.pdf>.

²¹ UNESCO: *Report of the World Commission on the Ethics of Scientific Knowledge and Technology on Robotics Ethics*, September 2017, SHS/YES/COMEST-10/17/2 REV, para. 38.

²² FRANKISH, K., RAMSEY, W. M. (eds.): *The Cambridge Handbook on Artificial Intelligence*, Cambridge: Cambridge University Press, 2014, p. 7.

namely systems or artefacts capable of performing tasks associated with human cognition, such as problem-solving, learning, or decision-making.²³

Against this background, the definitional ambiguity surrounding AI has emerged as one of the earliest and most persistent challenges in both policy and legal discourse. Early attempts to regulate AI, whether through ethical guidelines or legal instruments adopted by States, international organisations, or private actors, were characterised by a considerable degree of fragmentation. Individual instruments tended to adopt their own definitions of AI, often reflecting specific regulatory objectives, technological understandings, or institutional perspectives. This lack of conceptual convergence not only complicated international cooperation but also posed significant obstacles to the development of coherent regulatory frameworks, as legal regulation cannot effectively operate in relation to a concept that lacks sufficient definitional clarity.²⁴

Over time, however, a gradual convergence in the conceptualisation of AI can be observed. From a legal perspective, this trend is reflected in the definitions contained in the first binding international instruments addressing AI, most notably the *Artificial Intelligence Act* and the *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, as well as in a growing body of soft law instruments adopted since 2024. These developments indicate an emerging alignment in the understanding of AI as a distinct object of legal regulation. On this basis, an AI system may be defined as “a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.”²⁵ This definition reflects a broad, function-oriented understanding of AI systems, which

²³ COPELAND, J.: *Artificial Intelligence: Philosophical Introduction*, New Jersey: Wiley-Blackwell, 1993, pp. 1 and 43.

²⁴ KLUČKA, J.: *General Overview of the Artificial Intelligence and International Law*. In: KLUČKA, J., BAKOŠOVÁ, L., SISÁK, E. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Leges, 2021, p. 13.

²⁵ *AI Act*, Art. 3 (1). See also *Framework Convention*, Art. 2; OECD: *Explanatory Memorandum on the Updated OECD Definition of an AI System*. OECD Artificial Intelligence Papers, No. 8, 2024, p. 4. Online: https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf.

is particularly suited to legal regulation due to its flexibility and technological neutrality, allowing it to accommodate both existing and future technological developments.

Against this conceptual background, it becomes necessary to turn to the nature and structure of IHRL as the primary normative framework through which the implications of AI systems are to be systematically assessed.

2. THE EMERGENCE AND NATURE OF INTERNATIONAL HUMAN RIGHTS LAW

As AI increasingly reshapes social, economic, and governance structures, IHRL provides a critical normative framework through which the implications of these technologies can be assessed, ensuring their alignment with the principles of human dignity, equality, and freedom. In this regard, IHRL constitutes a specialised branch of public international law concerned with the overarching mission to protect universal features of the human being from the exercise of sovereign power.²⁶ Unlike classical public international law, which traditionally governed relations between sovereign states, IHRL is characterised by its individual-centred orientation, recognising individuals not merely as objects but as subjects of international law.²⁷ This shift is reflected not only in the substantive obligations imposed on States, but also in the development of institutional and procedural mechanisms designed to monitor compliance, provide remedies, and ensure the effective realisation of human rights.²⁸

From a historical perspective, some of the earliest foundations of human rights may be identified in dominant religious teachings, which, although not formulating human rights in their modern sense, emphasised moral duties and responsibilities towards others. At the same time, principles such as equality, human dignity, and the sanctity of human life, embedded within these traditions, later became important normative foundations for the development of human rights concepts.²⁹ These ideas were subsequently fur-

²⁶ MACKLEM, P.: *What is International Human Rights Law? Three Applications of a Distributive Account*. In: McGill Law Journal, Vol. 52, 2007, p. 577; DONNELLY, J.: *Universal Human Rights in Theory and Practice*, Second edition, Ithaca: Cornell University Press, 2003.

²⁷ TYMOFEYEVA, A.: *Jednotlivec jako subjekt mezinárodního práva: Příklad EÚLP*. In: ONDŘEJ, J.: *Postavení jednotlivce v mezinárodním právu: Proměny (vývoj) právního postavení jednotlivce v mezinárodním právu*. Praha: Univerzita Karlova, Právnická fakulta, 2020, p. 12; ELBERT, L.: *Rule of Law a mezinárodní právo*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2015, p. 212.

²⁸ KEITNER, CH.: *International Law Frameworks (Concepts and Insights)*. Fifth Edition, St. Paul: Foundation Press, 2021, p. 149.

²⁹ SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, p. 2. Online: https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2045&context=faculty_publications.

ther shaped through philosophical and legal developments, including natural law theories, constitutional traditions, humanitarian movements, and reactions to slavery, colonialism, and armed conflicts,³⁰ all of which contributed to the gradual internationalisation of human rights protection.³¹ Nevertheless, it was only the unprecedented scale of atrocities committed during the Second World War that fundamentally transformed the relationship between the individual and the international legal order and accelerated the emergence of international mechanisms for the protection of human rights.³²

The key milestone in the development of IHRL is the adoption of the *Universal Declaration of Human Rights* (hereinafter “UDHR”) in 1948.³³ Although the UDHR technically is not legally binding on States, its adoption marked the formal genesis of a profound structural transformation of the international legal order.³⁴ Building upon the UDHR, the *International Covenant on Civil and Political Rights* and the *International Covenant on Economic, Social and Cultural Rights* provide a more detailed codification of fundamental rights, collectively forming, together with the UDHR, the core of what is commonly referred to as the International Bill of Human Rights.

Rights protected within the International Bill of Human Rights include, in particular, the right to life, the prohibition of torture and slavery, freedom of expression, freedom of religion, the right to privacy, the right to a fair trial, political participation rights, as well as economic, social, and cultural rights, including the rights to work, education, health, and an adequate standard of living. Collectively, these rights form the normative foundation of the contemporary IHRL and continue to shape the interpretation and development of human rights protection at both the universal and regional levels.

Beyond these foundational instruments, the UN human rights framework has been further expanded through a wide range of specialised treaties, dec-

³⁰ KOVÁCS, P.: *Introductory Thoughts About the Universal Protection of Human Rights*. In: KOVÁCS, P., BERES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, pp. 15–16.

³¹ ŠMIGOVÁ, K.: *The History of Human Rights Law: Generations of Human Rights*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, p. 34.

³² HANSUNGULE, M.: *The Historical Development of International Human Rights*. In: CHOWDHURY, A. R., BHUIYAN, J. H. (eds.): *An Introduction to International Human Rights Law*, Brill: Nijhoff, 2010, p. 1.

³³ UN General Assembly: *International Bill of Human Rights: A Universal Declaration of Human Rights*, A/RES/217(III), 10 December 1948.

³⁴ MACKLEM, P.: *What is International Human Rights Law? Three Applications of a Distributive Account*. In: *McGill Law Journal*, Vol. 52, 2007, p. 578.

larations, and resolutions addressing specific categories of rights or vulnerable groups. In parallel, regional human rights systems developed in Europe, the Americas, and Africa, contribute to the pluralisation and strengthening of human rights protection.³⁵ These instruments contribute to the progressive development and refinement of human rights norms, reflecting both the increasing complexity of the international legal order and the evolving understanding of human rights in response to new societal and technological challenges.³⁶ At the same time, their adoption illustrates the gradual transformation of IHRL from a framework based primarily on declaratory principles into a more institutionalised system of legally binding obligations accompanied by monitoring and adjudicatory mechanisms operating at both the universal and regional levels.

The development of IHRL is often analysed through the lens of successive “generations” of human rights, which reflect the gradual broadening of the scope of protection over time. This tripartite classification is commonly associated with the ideals of the French Revolution—*liberté, égalité, and fraternité*—corresponding respectively to civil and political rights (first generation), economic, social and cultural rights (second generation), and collective or solidarity rights (third generation). The eighteenth and nineteenth centuries were marked by struggles aimed primarily at limiting authoritarian power and securing fundamental rights and freedoms, including the right to life, freedom of expression, association, religion, and political participation.³⁷ Civil and political rights are traditionally associated with obligations requiring States primarily to refrain from unjustified interference with individual freedoms. As the role of the state evolved in response to industrialisation and increasing social inequalities, greater emphasis was placed on the recognition of socio-economic rights, including the right to work, the right to education, and the right to an adequate standard of living. By contrast, the effective realisation of socio-economic rights generally requires active measures on the part of the State.³⁸ In more recent decades, processes of globalisation and a growing awareness of shared global challenges—particularly those

³⁵ See Chapter 3. *Normative Framework of International Human Rights Law*.

³⁶ STEINER, H. J., ALSTON, P.: *International Human Rights in Context: Law, Politics, Morals*. Second Edition, Oxford: Oxford University Press, 2000, pp. 140–141.

³⁷ VILJOEN, F.: *International Human Rights Law: A Short History*. In: UN Chronicle, No. 1 & 2, 2009, p. 8. Online: <https://www.un-ilibrary.org/content/journals/15643913/46/2/3/read>.

³⁸ ŠMIGOVÁ, K.: *The History of Human Rights Law: Generations of Human Rights*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, p. 49.

linked to poverty and development—have contributed to the emergence of so-called third generation rights, such as the right to a healthy environment, the right to self-determination, and the right to development.³⁹ These rights reflect the idea that individual rights cannot be fully realised without addressing broader social, environmental, and economic conditions that affect entire communities and nations. Solidarity rights are considered as “collective rights” because they pertain not just to individuals, but to groups of people or even to humanity as a whole.⁴⁰

Against this background, contemporary technological developments increasingly give rise to discussions concerning the possible emergence of a new generation of so-called fourth-generation rights.⁴¹ Much like earlier generations of rights emerged in response to broader societal, political, and economic transformations, current debates reflect the growing influence of digital technologies and AI on the enjoyment of human rights. While the first three generations of human rights were historically grounded in broader normative values such as liberty, equality, and solidarity, contemporary discussions concerning fourth-generation rights increasingly seek to identify principles capable of fulfilling a similar role within the digital environment and in the context of rapid technological and scientific development. At the same time, identifying a coherent normative foundation for these emerging rights remains particularly challenging due to the rapidly evolving nature of digital technologies and the risk of reducing such rights to fragmented or purely reactive responses to individual technological developments.⁴² From this perspective, the emerging fourth generation of human rights may be understood as seeking to ensure that individuals and their rights remain effectively protected in the digital age and that new technologies are developed

³⁹ VILJOEN, F.: *International Human Rights Law: A Short History*. In: UN Chronicle, No. 1 & 2, 2009, p. 8. Online: <https://www.un-ilibrary.org/content/journals/15643913/46/2/3/read>.

⁴⁰ ŠMIGOVÁ, K.: *The History of Human Rights Law: Generations of Human Rights*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, p. 51.

⁴¹ See e.g. BARONI, M. J. L.: *Fourth Generation Human Rights in View of the Fourth Industrial Revolution*. In: *Philosophies*, Vol. 9, No. 2, 2024, pp. 39–54; PEREPOLKIN, S et al.: *The concept of the fourth generation of human rights: fact or perspective of scientific discourse*. In: *Brazilian Journal of International Law*, Vol. 21, No. 2, 2024, pp. 280–295; SONG, L., MA, CH.: *Identifying the fourth generation of human rights in digital era*. In: *International Journal of Legal Discourse*, Vol. 7, No. 1, 2022, pp. 83–111.

⁴² BARONI, M. J. L.: *Fourth Generation Human Rights in View of the Fourth Industrial Revolution*. In: *Philosophies*, Vol. 9, No. 2, 2024, p. 43.

and used in a manner consistent with human dignity, including respect for liberty, equality, and solidarity within the human community.⁴³

In this regard, particular importance must be attached to the principle of human dignity, which constitutes one of the central normative foundations of contemporary IHRL and increasingly serves as a key interpretative framework for assessing the implications of AI and other emerging technologies for the protection of fundamental rights.⁴⁴ While early understandings of dignity were linked to social status and honour, later philosophical traditions increasingly emphasised the inherent worth and autonomy of every individual, particularly in Enlightenment thought and Kant's philosophy. Following the atrocities of the Second World War, human dignity became firmly embedded within contemporary IHRL.⁴⁵ Today, it is widely understood both as the foundation of human rights and as an interpretative principle guiding their development and application.⁴⁶ This understanding is reflected in core IHRL instruments, including the UDHR, the International Covenant on Civil and Political Rights, the International Covenant on Economic, Social and Cultural Rights,⁴⁷ as well as in regional human rights frameworks.⁴⁸ As for the functions it plays in the IHRL, Carozza identifies three principal functions. First, human dignity serves to support expansive interpretations of human rights by linking individual rights to a broader and universal value framework, thereby counterbalancing restrictive interpretations. Second, it strengthens the centrality and importance of specific rights while limiting

⁴³ ŠMIGOVÁ, K.: *The History of Human Rights Law: Generations of Human Rights*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, pp. 52–53.

⁴⁴ See e.g. NAGY, N.: "Humanity's new frontier": *Human rights implications of artificial intelligence and new technologies*. In: *Hungarian Journal of Legal Studies*, Vol. 64, No. 2, 2023, pp. 239–240.

⁴⁵ MCCRUDDEN, CH.: *Human Dignity and Judicial Interpretation of Human Rights*. In: *The European Journal of International Law*, Vol. 19, No. 4, 2008, pp. 656–663.

⁴⁶ TEO, S.A.: *Human dignity and AI: mapping the contours and utility of human dignity in addressing challenges presented by AI*. In: *Law, Innovation and Technology*, Vol. 15, 2023, p. 243.

⁴⁷ *International Covenant on Civil and Political Rights*, New York, 16 December 1966, Art. 10; *International Covenant on Economic, Social and Cultural Rights*, New York, 16 December 1966, Art. 13.

⁴⁸ *Protocol No. 13 to the Convention for the Protection of Human Rights and Fundamental Freedoms concerning the abolition of the death penalty in all circumstances*, Vilnius, 3 May 2002, ETS No. 187, Preamble; *African Charter on Human and Peoples' Rights*, Banjul, 27 June 1981, Art. 5; *American Convention on Human Rights: "Pact of San José, Costa Rica"*, San José, 22 November 1969, Art. 11 (1).

the scope for exceptions or restrictions. Third, human dignity functions as a broader normative reference point that extends beyond the literal wording of treaties and enables the development of treaty norms through interpretation. At the same time, the absence of a universally accepted understanding of dignity within international practice continues to raise interpretative challenges concerning its precise normative scope and application.⁴⁹

Contemporary technological developments therefore challenge not only the practical protection of human rights, but also the conceptual boundaries and normative foundations upon which IHRL has traditionally been constructed. Yet, the evolutionary character of IHRL also illustrates its capacity to adapt to changing societal conditions and emerging technological developments through the interpretation and progressive development of existing norms.

⁴⁹ CAROZZA, P. G.: *Human Dignity*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, p. 358.

3. NORMATIVE FRAMEWORK OF INTERNATIONAL HUMAN RIGHTS LAW

The growing use of AI systems raises the fundamental question of which normative standards should govern their development and deployment from the perspective of IHRL. Although contemporary AI technologies create novel forms of interference with fundamental rights, these challenges continue to be addressed primarily through the existing normative framework of IHRL, composed of international treaties, customary international law, general principles of law, and the interpretative practice of international and regional human rights bodies.⁵⁰ At the same time, the rapid pace of technological development has increased the importance of soft law instruments and interpretative mechanisms, which play a significant role in clarifying the application of human rights norms to emerging technological contexts. This Chapter therefore identifies the principal universal and regional human rights instruments forming the normative framework for assessing the human rights implications of AI systems and considers their capacity to respond to harms that are cross-sectoral, data-driven and frequently attributable to complex combinations of public and private actors.

3.1. Universal Normative Framework of Human Rights Protection

The universal human rights framework developed through the gradual internationalisation and legal consolidation of human rights protection following the adoption of the Charter of the United Nations in 1945. Although the UN Charter did not establish a comprehensive catalogue of human rights, it af-

⁵⁰ For instance, International Court of Justice, European Court of Human Rights, Inter-American Court of Human Rights, African Court on Human and Peoples' Rights, Human Rights Committee, Committee on Economic, Social and Cultural Rights, Committee on the Rights of the Child, African Commission on Human and Peoples' Rights, or the Inter-American Commission on Human Rights.

firmed the commitment of the international community to promoting universal respect for human rights and fundamental freedoms.⁵¹

The Universal Declaration of Human Rights, adopted in 1948, subsequently established the principal normative catalogue of civil, political, economic, social and cultural rights and provided the foundation for the development of binding international human rights treaties. Although formally non-binding, the Declaration has exercised considerable influence on international treaty-making, domestic constitutional law and international human rights practice, while parts of its content are widely regarded as reflecting customary international law.⁵²

The core of the universal treaty framework is formed by the International Bill of Human Rights, comprising the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights and the International Covenant on Economic, Social and Cultural Rights. Together, these instruments protect the principal rights relevant to the assessment of AI systems, including the rights to life, privacy, equality and non-discrimination, freedom of expression, fair trial, work, health, education, social security and an adequate standard of living.⁵³ While the ICCPR primarily establishes immediately applicable civil and political rights, several obligations under the ICESCR are subject to progressive realisation within the limits of available resources.⁵⁴

Beyond the two Covenants, the universal human rights framework has been further developed through a number of specialised treaties addressing specific forms of discrimination, particular categories of rights, and the protection of vulnerable groups, including women, children, persons with disabilities and racial or ethnic minorities.⁵⁵ Although formally separate le-

⁵¹ See Arts. 1 (3), 13 (1) (b), 55 (c), 56, 62 (2), 68 and 76 (c) of the Charter of the United Nations, San Francisco, 26 June 1945, 1 UNTS XVI.

⁵² E.g. HANNUM, H.: *The Status of the Universal Declaration of Human Rights in National and International Law*. In: Georgia Journal of International and Comparative Law, Vol. 25, 1996, p. 289; MERON, T.: *Human Rights and Humanitarian Norms as Customary Law*. Oxford: Clarendon Press, 1989.

⁵³ *International Covenant on Civil and Political Rights*, New York, 16 December 1966, Arts. 6–27; *International Covenant on Economic, Social and Cultural Rights*, New York, 16 December 1966, Arts. 6–15.

⁵⁴ *International Covenant on Economic, Social and Cultural Rights*, New York, 16 December 1966, Art. 2 (1)

⁵⁵ *International Convention on the Elimination of All Forms of Racial Discrimination*. New York, 21 December 1965; *Convention on the Rights of the Child*. New York, 20 November 1989; *Convention on the Elimination of All Forms of Discrimination against Women*. New

gal instruments, universal human rights treaties form part of a broader and interconnected system of international human rights protection linked by common principles, particularly human dignity, equality, non-discrimination, participation, accountability, and the effective protection of individuals against human rights violations.

This interconnected character is particularly important in the context of AI systems, whose impacts frequently affect multiple rights simultaneously and may disproportionately interfere with the rights of vulnerable individuals and groups. Rights such as the right to privacy, freedom of expression, equality and non-discrimination, the right to a fair trial, and economic and social rights, including the rights to work, health, education, and social security, are increasingly affected by the development and deployment of AI technologies. The cross-sectoral and data-driven nature of AI systems therefore requires their assessment against the universal human rights framework as a whole, rather than through isolated consideration of individual rights.

In addition to binding human rights treaties, soft law instruments also play an important role in clarifying the application of human rights standards to contemporary technological and economic realities. Particular attention should be paid to the *UN Guiding Principles on Business and Human Rights*, endorsed by the Human Rights Council in 2011. Although international human rights obligations remain primarily incumbent upon States, the growing role of private actors in the development, deployment, and operation of AI systems makes the business and human rights framework especially relevant. The UN Guiding Principles rest on three pillars: the State duty to protect human rights, the corporate responsibility to respect human rights, and the need to ensure access to effective remedies.⁵⁶ In the context of AI, these principles require companies to avoid causing or contributing to adverse human rights impacts, to address impacts directly linked to their operations, products or services, and to conduct human rights due diligence capable of identifying, preventing, mitigating, and accounting for AI-related risks. They also underline the importance of meaningful stakeholder engage-

York, 18 December 1979; *Convention on the Rights of Persons with Disabilities*, New York, 13 December 2006.

⁵⁶ UN Office of the High Commissioner for Human Rights: *Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework*, New York and Geneva, 2011, Principles 13–31. Online: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf.

ment, continuous monitoring, transparency, and accessible remedial mechanisms where AI systems cause or contribute to human rights harm.

3.2. Regional Human Rights Framework

Although it is recognised that a number of regional human rights systems exist worldwide, in light of the research focus of this monograph, attention will be devoted only to a brief overview of the legal frameworks developed within the COE, the EU, the AU and the OAS, as the role and practice of these organisations in the field of human rights protection are examined in greater detail in subsequent parts of the monograph.

3.2.1. European Human Rights Framework

In the European context, two interconnected but institutionally distinct systems of human rights protection have developed within the framework of the COE and the EU.

The Council of Europe constitutes the principal regional human rights system in Europe. Its establishment in 1949 formed part of broader post-war efforts to prevent renewed conflict, resist authoritarianism, and promote a common framework of democratic values and fundamental freedoms,⁵⁷ thereby placing the protection of human rights, democracy and the rule of law among the organisation's core objectives.⁵⁸

The central normative instrument of the COE human rights system is the *Convention for the Protection of Human Rights and Fundamental Freedoms* (also known as the “European Convention on Human Rights”), which established a binding catalogue of civil and political rights together with an international supervisory mechanism centred around the European Court of Human Rights.⁵⁹ The Convention is complemented by a number of additional protocols expanding both the catalogue of protected rights and procedural guarantees. The rights protected under the ECHR may be broadly classified

⁵⁷ HEYNS, CH., KILLANDER, M.: *Universality and the Growth of Regional Systems*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, p. 675.

⁵⁸ *Statute of the Council of Europe*, London, 5 May 1949, ETS No. 001, Art. 1 (b); SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, p. 16.

⁵⁹ *European Convention on Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11, 14 and 15*, Rome, 4 October 1950, ETS No. 005, Art. 2–14.

into three substantive categories: those relating to the protection of human life and personal integrity, those safeguarding the rule of law, and those underpinning pluralism and tolerance in democratic society.⁶⁰

The Convention is complemented by a broader network of Council of Europe treaties addressing social rights, the prevention of torture, minority protection and other specific areas of human rights protection.⁶¹ Together, these instruments constitute a multilayered normative framework⁶² capable of responding, through interpretation, to changing societal and technological conditions.

Although the ECHR was adopted long before the emergence of contemporary digital technologies, several of its guarantees are particularly relevant to AI, including the right to respect for private and family life, freedom of expression, the right to a fair trial and the prohibition of discrimination. Their technologically neutral formulation enables their application to new forms of interference arising from surveillance, biometric identification, algorithmic decision-making, profiling and other data-driven practices. The relevance of these guarantees therefore depends not on the creation of AI-specific rights within the Convention system, but primarily on their evolutionary interpretation in response to technologically mediated forms of interference.

The development of human rights protection within the framework of the EU followed a different trajectory. Originally established primarily as an economic integration project, the European Communities did not initially contain a comprehensive catalogue of fundamental rights. Nevertheless, through the jurisprudence of the Court of Justice of the European Union, fundamental rights gradually emerged as general principles of EU law, drawing inspiration from the constitutional traditions common to the Member States and from international human rights instruments, particularly the ECHR.⁶³ A major milestone in the development of the EU human rights framework was the adoption of the *Charter of Fundamental Rights of the European Union*,

⁶⁰ JANKUV, J.: *Medzinárodné a európske mechanizmy ochrany ľudských práv*. Bratislava: IURA EDITION, 2006, p. 118.

⁶¹ *European Social Charter*, Turin, 18 October 1961, ETS No. 35; *European Convention for the Prevention of Torture*, Strasbourg, 26 November 1987, ETS No. 126; *European Charter for Regional or Minority Languages*, Strasbourg, 5 November 1992, ETS No. 148; *Framework Convention for the Protection of National Minorities*, 1 February 1995, ETS No. 15.

⁶² SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, pp. 17–18.

⁶³ JANKUV, J.: *Medzinárodné a európske mechanizmy ochrany ľudských práv*. Bratislava: IURA EDITION, 2006, pp. 201–205.

proclaimed in 2000 and acquiring binding legal force with the entry into force of the *Treaty of Lisbon* in 2009. Pursuant to Art. 6 (1) Treaty on the European Union, the Charter has the same legal value as the Treaties⁶⁴ and therefore forms part of the primary law of the EU. Unlike traditional international human rights treaties, however, the Charter was not conceived as a classical international agreement concluded between States, but rather as an instrument integrated into the autonomous legal order of the EU.⁶⁵

The Charter consolidates a broad catalogue of civil, political, economic and social rights, structured around the values of dignity, freedoms, equality, solidarity, citizens' rights and justice. In the context of AI, several Charter rights are of particular relevance, including human dignity, the right to integrity of the person, respect for private and family life, freedom of expression and information, equality before the law, non-discrimination, the rights of the child, the rights of persons with disabilities, the right to an effective remedy and to a fair trial, and the right to good administration. Particularly significant is the express recognition of the right to the protection of personal data as an autonomous fundamental right. Pursuant to Art. 8 of the Charter, everyone has the right to the protection of personal data concerning him or her. Such data must be processed fairly, for specified purposes, and on the basis of the consent of the person concerned or another legitimate basis laid down by law. The Charter further guarantees the right of access to collected data, the right to rectification, and requires compliance with these rules to be subject to control by an independent authority.⁶⁶ Given its status as primary EU law, the Charter therefore provides a binding standard for EU institutions and for Member States when they act within the scope of EU law. This gives the Charter particular relevance in areas shaped by EU regulation, including data protection, digital services, automated decision-making and AI. The autonomous protection of personal data distinguishes the EU framework from the ECHR system, where data protection has traditionally developed primarily through the interpretation of the right to respect for private life. This distinction is particularly significant in the context of AI, whose operation frequently depends on extensive data collection, profiling, inference and automated processing. The coexistence of the ECHR and EU Charter

⁶⁴ *Treaty on the European Union and Treaty on the Functioning of European Union.*

⁶⁵ MAZÁK, J., JÁNOŠÍKOVÁ, M. et al.: *The Charter of Fundamental Rights of the European Union in Proceedings before Courts of the Slovak Republic*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2016, pp. 26–27.

⁶⁶ *Charter of Fundamental Rights of the European Union*, 26.10.2012, OJ C 326, Art. 8.

therefore provides a particularly dense regional framework for addressing AI-related human rights risks, combining technologically adaptable human rights guarantees with more specific protection of data-driven forms of interference.

3.2.2. African Human Rights Framework

The African regional human rights system is built on the *African Charter on Human and Peoples' Rights*. The Charter reflected the specific historical, political and social context of the African continent, such as post-colonial experiences and the struggle against apartheid and authoritarian regimes. The African Charter is often regarded as a distinctive regional instrument because it combines individual rights with peoples' rights and individual duties. In addition to civil and political rights, it protects collective and solidarity rights, including the rights to self-determination, development, peace, control over natural resources, and the right to a satisfactory environment.⁶⁷ In this respect, the African system adopted a broader and more collective understanding of human rights protection than other regional systems.⁶⁸ Unlike more technocratic rights-based frameworks emerging in other jurisdictions, the African human rights system is characterised by a broader communitarian vision that situates human dignity, equality, and development within a collective conception of social justice and shared societal responsibility.⁶⁹ The African system has subsequently been complemented by specialised instruments, including the *Protocol on the Rights of Women in Africa*⁷⁰ and the *African Charter on the Rights and Welfare of the Child*.⁷¹

In the context of AI and digital technologies, particular relevance may be attached to equality and non-discrimination, human dignity, fair trial guar-

⁶⁷ *African Charter on Human and Peoples' Rights*, Banjul, 27 June 1981, Art. 22–24.

⁶⁸ SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, p. 21; POŠIVÁKOVÁ, L.: Úloha regionalizmu v oblasti ochrany ľudských práv. In: GIERTL, A. (ed.): *Regionalizmus: Stav, východiská, perspektívy*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2013, p. 120.

⁶⁹ FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: *African Journal of Legal Research*, Vol. 4, No. 2, 2026, p. 61.

⁷⁰ *Protocol to the African Charter on Human and Peoples' Rights on the Rights of Women in Africa*, Maputo, 13 September 2000.

⁷¹ *African Charter on the Rights and Welfare of the Child*, OAU Doc. CAB/LEG/24.9/49 (1990), 11 July 1990.

antees, and freedom of expression and access to information.⁷² The collective and developmental dimensions of the African Charter may also be significant where AI systems generate broader societal effects, including unequal access to technological benefits, structural discrimination or risks of deepening existing socio-economic inequalities. Unlike several other regional human rights instruments, the African Charter does not expressly recognise a standalone right to privacy. Privacy-related protection has nevertheless been derived through the interpretation of related guarantees, particularly human dignity and personal security.⁷³ This interpretative development is especially relevant in the context of AI systems based on extensive data collection, profiling and surveillance, as it illustrates both the adaptability of the existing framework and the possible limits of protection where technologically specific guarantees remain underdeveloped.

3.2.3. *Inter-American Human Rights Framework*

The Inter-American human rights system developed within the framework of the OAS, which has recognised the protection of human rights as one of its core objectives. The foundations of the system may be traced to the adoption of the *American Declaration of the Rights and Duties of Man* in 1948.⁷⁴ The Declaration recognises a broad range of civil, political, economic, social, and cultural rights, including the rights to life, liberty, equality before the law, freedom of expression, religion, fair trial, education, work, and social security.⁷⁵ In addition to individual rights, the instrument also emphasises corresponding duties owed by individuals toward society and the community, such as the duty of the individual to conduct himself in relation to others that each and every one may fully form and develop his personality, or the duty to acquire at least an elementary education.⁷⁶ Although formally non-binding, the Declaration has acquired considerable legal and interpretative

⁷² *Ibid*, Art. 2–3, Art. 5, Art. 7, Art. 9.

⁷³ FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: *African Journal of Legal Research*, Vol. 4, No. 2, 2026, p. 62.

⁷⁴ As one of the earliest regional human rights instruments, the Declaration played a formative role in shaping the drafting of the Universal Declaration of Human Rights at the universal level.

⁷⁵ *American Declaration of the Rights and Duties of Man*, Ninth International Conference of American States, Bogotá, 1948, Chapter I.

⁷⁶ *Ibid*, Art. XXIX and XXXI.

significance within the Inter-American human rights framework and has influenced the subsequent development of regional human rights protection.

The regional framework was further strengthened by the adoption of the *American Convention on Human Rights* in 1969. The Convention protects core civil and political rights, including the rights to life, liberty, privacy, fair trial, freedom of expression, religion, assembly and association, political participation, equality before the law, and judicial protection, while socio-economic rights were initially addressed only in general terms through the concept of progressive realization.⁷⁷ More extensive protection of economic, social and cultural rights was introduced later through the *Protocol of San Salvador*, which expressly recognises rights such as the rights to health, education, work, social security, adequate food, and a healthy environment.⁷⁸

The Inter-American system has gradually expanded through the adoption of additional regional treaties dealing with specific human rights issues, including torture, forced disappearance, violence against women and the rights of persons with disabilities.⁷⁹ Some of these instruments represented pioneering international efforts and later inspired developments at the universal level.

In the context of AI, particular relevance attaches to the rights to privacy, equality and non-discrimination, freedom of expression, due process and judicial protection. These guarantees provide an important legal basis for addressing AI-related concerns involving surveillance, automated profiling, discriminatory decision-making, data processing and access to effective remedies. At the same time, the evolution of the Inter-American system demonstrates the capacity of regional human rights law to develop existing guarantees in response to new forms of interference rather than relying exclusively on the creation of technology-specific rights.

⁷⁷ *American Convention on Human Rights: "Pact of San José, Costa Rica"*, San José, 22 November 1969, Art. 26.

⁷⁸ *Additional Protocol to the American Convention on Human Rights in the Area of Economic, Social and Cultural Rights: Protocol of San Salvador*, San Salvador, 17 November 1988, Art. 6, 9–13.

⁷⁹ *Inter-American Convention to Prevent and Punish Torture*, Cartagena de Indias, 9 December 1985; *Inter-American Convention on Forced Disappearance of Persons*, Belém de Pará, 9 June 1994; *Inter-American Convention on the Prevention, Punishment and Eradication of Violence against Women*, Belém de Pará, 9 June 1994; *Inter-American Convention on the Elimination of All Forms of Discrimination against Persons with Disabilities*, Guatemala City, 7 June 1999.

4. INSTITUTIONAL FRAMEWORK FOR THE PROTECTION OF HUMAN RIGHTS

The effectiveness of IHRL depends not only on the existence of norms but also on the availability of mechanisms for their implementation and enforcement. The institutional architecture of IHRL combines judicial and quasi-judicial mechanisms with political, monitoring and advisory processes operating at both universal and regional levels. Rather than offering a comprehensive overview of all international human rights instruments, this Chapter outlines the core normative pillars of the contemporary human rights system that are particularly relevant for the subsequent analysis of AI and its implications for human rights protection.

At the universal level, attention is devoted to the UN human rights system, including the activities of the Human Rights Council (hereinafter “HRC”), the Office of the United Nations High Commissioner for Human Rights (hereinafter “OHCHR”), and the UN treaty bodies. From a regional perspective, attention is paid to the European systems developed within the framework of the COE and the EU, while also addressing regional human rights protection mechanisms operating under the AU and the OAS. The delineation of their functions and competences is particularly important in the context of invoking responsibility and ensuring accountability for human rights violations arising from the use of AI systems. The purpose of this Chapter is therefore not merely to describe the competences of existing human rights bodies, but to establish the institutional baseline against which their capacity to respond to AI-related harms can subsequently be assessed. This is particularly important because AI-related interferences may be transnational, technologically complex and attributable to multiple public and private actors, thereby challenging mechanisms originally designed primarily to address identifiable violations attributable to States.

4.1. The United Nations Human Rights Protection Mechanism

Within the universal framework of IHRL, the UN has developed a complex institutional system aimed at the promotion, monitoring and protection of human rights. At the centre of this system stands the HRC, the successor to the Human Rights Committee, whose aim is to promote universal

respect and protection of all human rights and fundamental freedoms,⁸⁰ the OHCHR, and the UN human rights treaty bodies, which perform complementary political, monitoring, interpretative and quasi-judicial functions.⁸¹

4.1.1. Charter-Based Mechanisms: Human Rights Council and the Office of the United Nations High Commissioner for Human Rights

The HRC acts as a political forum to discuss, address, decide, make recommendations, and report on all thematic human rights issues and situations throughout the world.⁸² This is done through the Universal Periodic Review, Special Procedures, complaints procedure, an Advisory Committee, and three Forums: on Minority Issues, Business and Human Rights, and Social Issues. For the purposes of AI-related human rights protection, these mechanisms are particularly relevant because they combine monitoring, standard-setting, thematic expertise and engagement with States and other stakeholders.

The Universal Periodic Review (UPR) constitutes the only universal mechanism through which the human rights situation of all UN Member States is periodically examined within the framework of the HRC. Based on principles of dialogue, cooperation, and equal treatment of States, the mechanism aims to improve human rights protection globally by enabling states to present measures undertaken to fulfil their human rights obligations and by facilitating the exchange of recommendations and best practices.⁸³ The UPR was designed as a cooperative and non-confrontational process intended to complement, rather than replace, the work of treaty bodies and Special Procedures. In addition to its monitoring function, the mechanism is closely linked to technical assistance and capacity-building initiatives aimed at sup-

⁸⁰ UN General Assembly: *Human Rights Council*, A/RES/60/251, 15 March 2006, para. 2.

⁸¹ The UN human rights treaty bodies collectively represent Committee on the Elimination of Racial Discrimination; Committee on Economic, Social and Cultural Rights; Human Rights Committee; Committee on the Elimination of Discrimination against Women; Committee against Torture; Committee on the Rights of the Child; Committee on Migrant Workers; Subcommittee on Prevention of Torture and other Cruel, Inhuman or Degrading Treatment or Punishment; Committee on the Rights of Persons with Disabilities; and Committee on Enforced Disappearances.

⁸² UN General Assembly: *Human Rights Council*, A/RES/60/251, 15 March 2006, para. 5. KOTHARI, M.: *From Commission to the Council: Evolution of UN Charter Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, p. 587.

⁸³ UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007, para. 4.

porting the implementation of recommendations. Its universal scope makes it potentially relevant for identifying systemic AI-related human rights concerns at the national level, including discriminatory automated systems, digital surveillance or insufficient regulatory safeguards. At the same time, concerns have been raised regarding the risk of the process becoming overly formalistic, particularly given the absence of enforcement mechanisms or sanctions for non-compliance. Despite these criticisms, the UPR remains an important platform for inclusive human rights dialogue and peer review at the universal level.⁸⁴

On the other side, the Special Procedures of the HRC constitute a system of independent human rights experts mandated to examine, monitor, report on, and provide guidance concerning thematic or country-specific human rights issues.⁸⁵ Frequently described as one of the most significant mechanisms within the UN human rights system, Special Procedures play an important role in the development of soft-law standards, the interpretation of international human rights obligations, and the promotion of accountability at both the national and international levels.⁸⁶ Their activities include country visits, communications with states, thematic reporting, and engagement with civil society, international organisations, and other stakeholders. Over time, the Special Procedures system has expanded considerably, reflecting both the growing complexity of global human rights challenges and the increasing politicisation of mandate creation and renewal processes within the HRC. While the proliferation of mandates has strengthened thematic coverage and visibility of human rights concerns, it has also generated challenges relating to institutional coherence, resource allocation, and the effectiveness of the system.⁸⁷ Despite these limitations, Special Procedures remain among the most visible and influential mechanisms within the UN human

⁸⁴ LAKATOS, I.: *Human Rights Protection in the UN: A General and Institutional Overview (The Former Commission on Human Rights, Human Rights Council, High Commissioner for Human Rights, the International Court of Justice and the Protection of Human Rights)*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, pp. 127–129.

⁸⁵ UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007, para. 39 and 45.

⁸⁶ LAKATOS, I.: *Human Rights Protection in the UN: A General and Institutional Overview (The Former Commission on Human Rights, Human Rights Council, High Commissioner for Human Rights, the International Court of Justice and the Protection of Human Rights)*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, p. 124.

⁸⁷ *Ibid.*, pp. 124–126.

rights framework, particularly in addressing emerging issues such as digital technologies, AI, surveillance, and data governance.⁸⁸ In the context of AI, this flexibility constitutes one of the principal strengths of the Special Procedures system, since it allows emerging risks to be addressed before they are reflected in treaty interpretation or judicial practice. Its weakness lies in the fact that such responsiveness is not accompanied by binding enforcement powers.

In accordance with the HRC Institutional Building Package, which introduced the current complaint procedure, individuals, groups, and non-governmental organisations are able to claim to be victims of human rights violations, or possessing reliable information about such violations, to submit complaints to the HRC.⁸⁹ Complaints undergo an initial admissibility review conducted by the Working Group on Communications and the Secretariat. Complaints may be submitted against any UN Member State, irrespective of whether the state has ratified a particular human rights treaty or entered reservations to specific instruments.⁹⁰ Manifestly unfounded or anonymous complaints are rejected, while admissible complaints are transmitted to the state concerned for observations. To be admissible, complaints must contain a factual description of the alleged violations, be compatible with the principles of the UN Charter and international human rights instruments, and generally require the exhaustion of domestic remedies unless such remedies

⁸⁸ See e.g. UN Human Rights Council: *Non-take-up of rights in the context of social protection: Report of the Special Rapporteur on extreme poverty and human rights*, Olivier De Schutter, A/HRC/RES/50/38, 19 April 2022; UN Human Rights Council: *Racial and xenophobic discrimination and the use of digital technologies in border and immigration enforcement: Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, E. Tendayi Achiume, A/HRC/RES/48/76, 17 December 2021; UN Human Rights Council: *Racial discrimination and emerging digital technologies: a human rights analysis: Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, A/HRC/RES/44/57, 18 June 2020; UN Human Rights Council: *Report of the Special Rapporteur on the right to privacy*, Joseph A. Cannataci, A/HRC/RES/34/60, 24 February 2018; UN Human Rights Council: *Rights of persons with disabilities: Report of the Special Rapporteur on the rights of persons with disabilities*, A/HRC/RES/49/52, 28 December 2021.

⁸⁹ UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007, para. 85.

⁹⁰ LAKATOS, I.: *Human Rights Protection in the UN: A General and Institutional Overview (The Former Commission on Human Rights, Human Rights Council, High Commissioner for Human Rights, the International Court of Justice and the Protection of Human Rights)*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection Of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, p. 129.

are ineffective or unreasonably prolonged.⁹¹ Once the above-mentioned requirements have been satisfied, the HRC examines the situation, generally in a confidential manner, and may decide to discontinue consideration, continue monitoring, appoint an independent expert, initiate public consideration of the matter, or recommend technical assistance and capacity-building measures for the state concerned.⁹² Importantly, the complaint procedure is not intended to function as an individual remedial mechanism and does not provide compensation to alleged victims. Its relevance to AI-related harms therefore lies primarily in its capacity to draw attention to systemic patterns of violations rather than to remedy individual algorithmic decisions.

Within the UN system, OHCHR occupies an important position in the protection of human rights and in shaping emerging approaches to the regulation of AI. It exercises a broad and multifaceted mandate aimed at promoting and ensuring the effective enjoyment of all human rights. Acting under the authority of the UN Secretary-General and within the mandate conferred by the UN General Assembly, the High Commissioner engages in constructive dialogue with States, provides advisory services and technical assistance upon request, and contributes to the development, interpretation, and implementation of international human rights standards. The OHCHR further performs a coordinating function across the UN system, supports the prevention of human rights violations, and facilitates international co-operation, while simultaneously addressing emerging normative and operational challenges,⁹³ including those arising from the rapid development and deployment of AI technologies. Its cross-cutting mandate is particularly relevant to AI governance because it enables human rights considerations to be integrated across different policy areas and institutional processes.

Taken together, the Charter-based mechanisms provide important avenues for identifying, interpreting and publicly addressing AI-related human rights risks. Their principal strength lies in universality, thematic flexibility and their capacity to react to emerging issues without requiring the creation of new treaty mechanisms. Their principal limitation, however, lies in the predominantly political or non-binding character of their outputs. They

⁹¹ UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007, para. 87. ELBERT, L.: *Rule of Law a medzinárodné právo*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2015, pp. 214–216.

⁹² UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007, para. 109.

⁹³ UN General Assembly: *High Commissioner for the promotion and protection of all human rights*, A/RES/48/141, 20 December 1993, para. 4.

therefore contribute primarily to standard-setting, monitoring and preventive oversight rather than to enforceable individual remedies.

4.1.2. Treaty-Based Mechanisms: UN Human Rights Treaty Bodies

Complementing the Charter-based mechanisms of the HRC and the OHCHR, the UN system further encompasses a treaty-based framework composed of specialised human rights treaty bodies established under the core international human rights conventions. Of particular relevance to AI-related human rights protection are the Human Rights Committee, the Committee on Economic, Social and Cultural Rights, the Committee on the Elimination of Racial Discrimination, the Committee on the Elimination of Discrimination against Women, the Committee on the Rights of the Child, and the Committee on the Rights of Persons with Disabilities. These bodies supervise state compliance with treaty obligations and contribute to the interpretation and progressive development of international human rights law in response to contemporary challenges,⁹⁴ including those arising from AI and digital technologies. The treaty bodies perform a range of supervisory and interpretative functions aimed at monitoring the implementation of international human rights obligations by States Parties. Their principal activities include the examination of periodic state reports,⁹⁵ the adoption of concluding observations and recommendations, the interpretation of treaty provisions through general comments, and, where applicable, the consideration of individual and interstate complaints.⁹⁶ Through these mechanisms, treaty bodies contribute significantly to the clarification, development, and practical application of international human rights law within the universal system.⁹⁷

Periodic reporting remains one of the principal mechanisms through which treaty bodies assess State implementation of human rights obligations. The process is informed not only by State reports, but also by information submitted by civil society organisations, national human rights institutions,

⁹⁴ NAGY, N.: “Humanity’s new frontier”: Human rights implications of artificial intelligence and new technologies. In: Hungarian Journal of Legal Studies, Vol. 64, No. 2, 2023, p. 260.

⁹⁵ E.g. *International Covenant on Civil and Political Rights*, New York, 16 December 1966, Art. 40.

⁹⁶ E.g. *Ibid*, Art. 41–42.

⁹⁷ See e.g. RODLEY, N. S.: *The Role and Impact of Treaty Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 621–648.

UN mechanisms and other actors. This broader evidentiary framework enables treaty bodies to assess the practical implementation of treaty obligations beyond formal legislative commitments presented by states.⁹⁸ In the context of AI, this enables treaty bodies to address systemic concerns such as discriminatory automated decision-making, digital exclusion, surveillance or insufficient protection of vulnerable groups even where no AI-specific treaty obligations exist.

Several treaties and optional protocols further establish procedures enabling treaty bodies to examine individual communications submitted by persons alleging violations of treaty rights.⁹⁹ Such procedures generally require the exhaustion of domestic remedies and are conducted through written and confidential proceedings involving both the complainant and the respondent state. Treaty bodies may also request interim measures where there is a risk of irreparable harm. The outcome of these proceedings typically takes the form of “Views” or decisions assessing whether the state concerned has violated its treaty obligations and recommending measures aimed at remedying the violation and preventing similar breaches in the future. Depending on the nature of the violation, the recommended measures may include judicial or administrative remedies, compensation, or the adoption of legislative and institutional reforms designed to prevent future human rights violations. Following the adoption of its Views, the Committees also monitor the implementation of the measures communicated to the respondent State.¹⁰⁰ Although the views are formally non-binding,¹⁰¹ they possess considerable interpretative authority and practical influence within the interna-

⁹⁸ RODLEY, N. S.: *The Role and Impact of Treaty Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 626–631.

⁹⁹ *Optional Protocol to the International Covenant on Civil and Political Rights*, New York, 16 December 1966; *Optional Protocol to the Convention on the Elimination of All Forms of Discrimination against Women*, New York, 6 October 1999; *Optional Protocol to the Convention on the Rights of Persons with Disabilities*, New York, 13 December 2006; *Optional Protocol to the International Covenant on Economic, Social and Cultural Rights*, New York, 10 December 2008.

¹⁰⁰ JANKUV, J.: *Medzinárodné a európske mechanizmy ochrany ľudských práv*. Bratislava: IURA EDITION, 2006, pp. 58–67 and 100–101; ELBERT, L.: *Rule of Law a medzinárodné právo*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2015, pp. 213–214, ŠTURMA, P.: *Mezinárodní a evropské kontrolní mechanismy v oblasti lidských práv*. 4. vydání, Praha: C.H. Beck, 2026, pp. 8–23.

¹⁰¹ E.g. *Optional Protocol to the International Covenant on Civil and Political Rights*, New York, 16 December 1966, Art. 5(4).

tional human rights system.¹⁰² Their potential relevance to AI lies in the possibility of applying established treaty guarantees to individual harms arising from automated decision-making, profiling, data processing or algorithmic discrimination. At the same time, the case-specific, reactive and procedurally demanding character of individual communications limits their capacity to address systemic or rapidly evolving technological risks.

Treaty bodies also contribute to the progressive development of international human rights law through the adoption of general comments, which clarify the meaning and scope of treaty provisions and adapt them to evolving social, legal, and technological realities.¹⁰³ This interpretative function illustrates one of the central strengths of IHRL in the context of AI: existing treaty norms can acquire new practical meaning without the need for formal treaty amendment. The drafting process usually involves several stages, including the appointment of a rapporteur, the preparation of draft texts and consultations with states, non-governmental organisations and other stakeholders. Although formally non-binding, they possess considerable normative authority and are widely relied upon by courts, governments and scholars.¹⁰⁴

Most human rights treaties also provide mechanisms for interstate complaints, enabling one state party to allege violations committed by another state party.¹⁰⁵ In practice, however, this mechanism has been used extremely

¹⁰² RODLEY, N. S.: *The Role and Impact of Treaty Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 634–636; JANKUV, J.: *Medzinárodné a európske mechanizmy ochrany ľudských práv*. Bratislava: IURA EDITION, 2006, pp. 58 and 100; ELBERT, L.: *Rule of Law a medzinárodné právo*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2015, pp. 213–214.

¹⁰³ See e.g. Committee on the Rights of the Child: *General Comment No. 25 (2021) on children's rights in relation to the digital environment*, (CRC/C/GC/25), 2 March 2021; Committee on the Rights of the Child: *General Comment No. 16 (2013) on State obligations regarding the impact of the business sector on children's rights*, CRC/C/GC/16, 17 April 2013; UN Human Rights Committee: *CCPR General Comment No. 16: Article 17 (Right to Privacy), The Right to Respect of Privacy, Family, Home and Correspondence, and Protection of Honour and Reputation*, 8 April 1988

¹⁰⁴ RODLEY, N. S.: *The Role and Impact of Treaty Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 631–632.

¹⁰⁵ *International Convention on the Elimination of All Forms of Racial Discrimination*, New York, 21 December 1965, Art. 11–13; *International Covenant on Civil and Political Rights*, New York, 16 December 1966, Art. 41–43; *Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*, New York, 10 December 1984, Art. 21; *International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families*, New York, 18 December 1990, Art. 74.

rarely within the UN treaty body system. States are generally reluctant to initiate contentious proceedings against one another in the sensitive field of human rights, partly because such actions may negatively affect diplomatic relations or expose the complaining state to counterclaims concerning its own conduct. Interstate complaint procedures have therefore remained largely dormant within the universal system.¹⁰⁶

Taken together, the treaty bodies demonstrate the capacity of the existing human rights framework to adapt to technological change through interpretation rather than formal normative replacement. Their principal strengths lie in specialised treaty expertise, the progressive interpretation of existing rights and the ability to identify both individual and systemic patterns of harm. Their effectiveness in the field of AI nevertheless remains constrained by the non-binding character of many outputs, limited enforcement mechanisms and the predominantly reactive nature of individual complaint procedures.

4.2. Regional Institutional Mechanisms for the Protection of Human Rights

The institutional and enforcement framework of IHRL is not limited to the universal system established under the auspices of the UN. Alongside global mechanisms, regional human rights systems play an increasingly important role in the interpretation, implementation and enforcement of human rights obligations. The most significant strengthening of human rights protection has taken place primarily within regional organisations, which have developed increasingly effective mechanisms for addressing human rights violations and providing remedies for resulting harm. Their principal contribution lies in the development of judicial and quasi-judicial mechanisms capable of providing more direct forms of redress to individuals whose rights have been violated. Due to their closer geographical, political and cultural proximity to Member States, regional mechanisms are often capable of responding more dynamically to emerging societal and technological challenges, including those associated with AI and digital technologies. They, in principle, operate as subsidiary systems of protection, providing avenues of redress where alleged violations of human rights have not been adequately

¹⁰⁶ RODLEY, N. S.: *The Role and Impact of Treaty Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, p. 633.

ly remedied at the domestic level. Within these systems, commissions and courts perform complementary functions. While courts primarily adjudicate contentious cases and may issue legally binding judgments, commissions often engage in monitoring activities, examine human rights situations within member states, process individual complaints, promote friendly settlements, and issue recommendations aimed at preventing future violations.¹⁰⁷

4.2.1. European Human Rights Protection Mechanisms

The European regional system of human rights protection is characterised by a particularly developed institutional and judicial framework operating within both the COE and the EU.

The supervisory mechanism of the ECHR is the European Court of Human Rights (hereinafter “ECtHR”), which was the first international court specifically established for the protection of human rights and one of the first systems enabling individuals to directly challenge violations of their rights at the international level.¹⁰⁸ The system is based primarily on inter-state and individual applications. The inter-state procedure allows any Contracting State to bring before the Court allegations that another Contracting State has violated the provisions of the Convention, without the need to demonstrate a specific interest or direct injury. By contrast, the individual application procedure enables individuals, groups of individuals, and non-governmental organisations to submit complaints against a Contracting State under whose jurisdiction they fell at the time of the alleged violation.¹⁰⁹ The Court may examine an application only after all available domestic remedies have been exhausted and the application has been submitted within four months from the final domestic decision. Applications are inadmissible if they are anonymous, substantially identical to cases already examined by the Court or another international body without presenting relevant new information, incompatible with the Convention, manifestly ill-founded, abusive, or where the applicant has not suffered a significant disadvantage. The Court may reject inadmis-

¹⁰⁷ POŠIVÁKOVÁ, L.: *Úloha regionalizmu v oblasti ochrany ľudských práv*. In: GIERTL, A. (ed.): *Regionalizmus: Stav, východiská, perspektívy*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2013, p. 121.

¹⁰⁸ SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, p. 17.

¹⁰⁹ JANKUV, J.: *Medzinárodné a európske mechanizmy ochrany ľudských práv*. Bratislava: IURA EDITION, 2006, pp. 121–122.

sible applications at any stage of the proceedings.¹¹⁰ Judgments of the Court are binding on the parties to the case, and States are obliged to comply with them. The execution of judgments is supervised by the Committee of Ministers, which may request the Court to clarify interpretative issues concerning a judgment. If a State refuses to comply with a final judgment, the Committee of Ministers may refer the matter back to the Court to determine whether the State has failed to fulfil its obligations under the Convention.¹¹¹ The Convention also grants the Court advisory jurisdiction. At the request of the Committee of Ministers, the Court may issue advisory opinions concerning the interpretation of the Convention and its Protocols. However, this competence does not extend to questions relating to the substantive scope of protected rights or matters connected to contentious proceedings. Requests for advisory opinions require a two-thirds majority within the Committee of Ministers, while the Court itself determines whether the request falls within its jurisdiction. Advisory opinions must be reasoned and may include separate opinions of individual judges.¹¹² This advisory function has been further expanded through *Protocol No. 16 to the Convention*, which entered into force in 2018. Under Protocol No. 16, the highest domestic courts and tribunals of the High Contracting Parties may request advisory opinions from the ECtHR on questions of principle relating to the interpretation or application of the rights and freedoms defined in the Convention and its Protocols. Unlike the general advisory jurisdiction under Article 47, Protocol No. 16 establishes a mechanism aimed at strengthening judicial dialogue between the Court and national courts and enhancing the consistent interpretation of the Convention within domestic legal systems. Although advisory opinions issued under Protocol No. 16 are not legally binding, they carry significant interpretative authority and contribute to the development of the Convention system and the harmonisation of human rights standards across Europe.¹¹³

From the perspective of AI, the significance of the ECtHR lies less in the existence of AI-specific jurisprudence than in the adaptability of established Convention standards. Principles developed in cases concerning surveillance, privacy, data retention, freedom of expression, procedural fairness and

¹¹⁰ *European Convention on Human Rights and Fundamental Freedoms*, Rome, 4 October 1950, Art. 35.

¹¹¹ *Ibid*, Art. 46.

¹¹² *Ibid*, Art. 47–49.

¹¹³ *Protocol No. 16 to the Convention on the Protection of Human Rights and Fundamental Freedoms*, Strasbourg, 2 October 2013, ETS No. 214, Art. 1 and 5.

non-discrimination provide an important basis for assessing AI-enabled interferences. At the same time, the Court remains predominantly reactive and dependent on individual applications arising after an alleged violation has occurred, which may limit its capacity to address systemic or preventive dimensions of AI governance.

Within the EU, the protection of fundamental rights is entrusted to the Court of Justice of the European Union (hereinafter “CJEU”) and the General Court of the European Union. Through actions for annulment, preliminary ruling proceedings, and infringement procedures, the EU courts contribute to the interpretation and enforcement of fundamental rights protected under the Charter of Fundamental Rights of the European Union.¹¹⁴ Individuals may, under certain conditions, challenge acts of EU institutions directly before EU courts and may seek compensation for damages resulting from unlawful conduct attributable to the Union. If the CJEU finds that the Member State has breached its obligations under EU law, including those arising from the Charter, the state concerned is required to take all necessary measures to bring the violation to an end and ensure compliance with fundamental rights standards protected within the EU legal order.¹¹⁵ The CJEU is particularly relevant to AI-related human rights protection because its jurisprudence has developed extensive standards concerning privacy, personal data, automated decision-making and effective judicial protection. Unlike the ECtHR, the CJEU operates within an integrated supranational legal order in which its judgments directly influence the interpretation and application of EU secondary legislation, including data protection and AI regulation. This institutional position enables the Court not only to remedy individual violations, but also to shape the regulatory conditions under which AI systems are developed and deployed across the Union.

The European framework therefore combines two complementary judicial models. The ECtHR provides external international supervision of State compliance with Convention rights, while the CJEU operates within the autonomous legal order of the EU and directly contributes to the interpretation of both fundamental rights and technology-specific regulation. Their

¹¹⁴ *Consolidated version of the Treaty on the Functioning of the European Union*, 2008/C 115/01, 13 December 2007, Art. 263, 265, 268.

¹¹⁵ BENKO, R.: *Právne prostriedky nápravy porušenia základných práv v európskom regióne v komparatívnej perspektíve s osobitným ohľadom na prostriedky nápravy finančnej povahy porušenia práv uznaných Chartou základných práv EÚ v horizontálnych vzťahoch*. In: GIERTL, A. (ed.): *Regionalizmus: Stav, východiská, perspektívy*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2013, p. 133.

coexistence creates a comparatively dense system of judicial protection, although differences in jurisdiction, applicable law and procedural access may also generate complexity and fragmentation. In the context of AI, this dual structure represents both a strength and a coordination challenge.

4.2.2. African Human Rights Protection Mechanisms

The African human rights system is primarily based on the African Charter on Human and Peoples' Rights and operates through the complementary functions of the African Commission on Human and Peoples' Rights (hereinafter "ACHPR") and the African Court on Human and Peoples' Rights. The Commission serves as the principal quasi-judicial body responsible for promoting, interpreting, and monitoring compliance with the Charter.¹¹⁶ Its primary focus is on the examination of complaints concerning alleged human rights violations. Such complaints may be submitted by individuals, non-governmental organisations, or States Parties to the African Charter claiming that another State Party has violated the rights guaranteed under the Charter. Under Art. 56 of the African Charter, individual communications submitted to the African Commission are admissible only if certain conditions are fulfilled. In particular, communications must identify their authors, be compatible with the Charter, avoid insulting language, not rely exclusively on media reports, and concern matters that have not already been settled internationally. Applicants must generally exhaust available domestic remedies and submit the communication within a reasonable period thereafter, unless domestic remedies are ineffective or unduly prolonged.¹¹⁷ Unlike the other systems, the African system envisages not only inter-state and individual communications procedures, but a special procedure for situations of gross and systematic violations.¹¹⁸ Although it may adopt recommendations, resolutions, and general comments, these instruments are generally non-binding and rely largely on voluntary state compliance.¹¹⁹ This represents an important structural limitation where human rights risks require timely and enforceable remedies.

¹¹⁶ *African Charter on Human and Peoples' Rights*, Banjul, 27 June 1981, Art. 30.

¹¹⁷ *Ibid.*, Art. 55–56.

¹¹⁸ SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, p. 21.

¹¹⁹ FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: *African Journal of Legal Research*, Vol. 4, No. 2, 2026, p. 66.

An important step towards strengthening human rights protection mechanism was the adoption of the Protocol to the African Charter on Human and Peoples' Rights on the Establishment of an African Court on Human and Peoples' Rights. Adopted in 1998 and entering into force in 2004, the Protocol established the African Court on Human and Peoples' Rights with the aim of reinforcing the protection and enforcement of human rights on the African continent.¹²⁰ The African Court provides a judicial mechanism for addressing violations of the Charter, including cases involving structural or collective harms. However, access to the Court remains relatively limited. While States and African intergovernmental organisations may submit cases directly, individuals and NGOs may access the Court only where the respondent State has made a special declaration under Art. 34 (6) of the Protocol.¹²¹ In practice, the limited number of such declarations, together with broader procedural and institutional constraints, has significantly restricted access to judicial protection within the African system.¹²² Furthermore, the Court may issue advisory opinions on legal matters relating to the African Charter or other relevant human rights instruments at the request of AU Member States, the AU, its organs, or recognised African organisations. However, the Court cannot provide an advisory opinion on matters that are already pending before the African Commission on Human and Peoples' Rights.¹²³

From the perspective of AI-related human rights protection, the African system therefore combines considerable normative potential with significant institutional constraints. The broad protection of individual and collective rights under the African Charter may accommodate harms arising from discrimination, surveillance, digital exclusion or unequal access to technological benefits. However, the limited enforceability of Commission outputs and restricted individual access to the Court may weaken the availability of effective

¹²⁰ JUMA, D.: *Access to the African Court on Human and Peoples' Rights: A Case of the Poacher turned Gamekeeper*. In: Essex Human Rights Review, Vol. 4, No. 2, 2007, p. 2.

¹²¹ *Protocol to the African Charter on Human and Peoples' Rights on the Establishment of an African Court on Human and Peoples' Rights*, Ouagadougou, 10 June 1998, Art. 34 (6).

¹²² FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: African Journal of Legal Research, Vol. 4, No. 2, 2026, pp. 66–67; LAZORČÁKOVÁ, T.: *Prístup mimovládnych organizácií a jednotlivcov k africkému systému ľudských práv*. In: STUDIA IURIDICA Cassoviensia, Vol. 5, No. 2, 2017, p. 16.

¹²³ *Protocol to the African Charter on Human and Peoples' Rights on the Establishment of an African Court on Human and Peoples' Rights*, Ouagadougou, 10 June 1998, Art. 5.

tive remedies where AI-related harms are dispersed, systemic or affect vulnerable groups.

4.2.3. Inter-American Human Rights Protection Mechanisms

Within the Inter-American human rights system, the Inter-American Court of Human Rights and the regional Commission on Human Rights perform complementary functions.

The Commission is primarily responsible for promoting and protecting human rights within the regional system. Its functions include promoting human rights awareness, issuing recommendations to Member States, preparing studies and reports, requesting information from governments on human rights matters, providing advisory assistance to States, examining individual petitions and communications concerning alleged human rights violations, and submitting annual reports to the General Assembly of the OAS.¹²⁴ Individual petitions may be submitted by persons, groups, or recognised non-governmental organisations alleging violations of Convention rights by a State Party. Their admissibility is subject to several conditions, including the exhaustion of domestic remedies, submission within six months of the final domestic decision, and the absence of parallel international proceedings concerning the same matter. Exceptions to the exhaustion requirement apply where domestic remedies are unavailable, ineffective, or subject to unreasonable delay. The Commission must declare petitions inadmissible if these requirements are not fulfilled, if the complaint is manifestly unfounded, or if the matter has already been examined by another international body.¹²⁵ The Convention also provides for interstate complaints, provided that the States concerned have expressly recognised the Commission's competence in this regard,¹²⁶ and to appoint special rapporteurs to prepare studies on hemisphere-wide problems.¹²⁷

The Inter-American Court of Human Rights constitutes the judicial component of the regional system and has jurisdiction over cases concerning the interpretation and application of the American Convention on Human

¹²⁴ *American Convention on Human Rights: "Pact of San José, Costa Rica"*, San José, 22 November 1969, Art. 41; ELBERT, L.: *Rule of Law a medzinárodné právo*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2015, p. 222.

¹²⁵ *Ibid*, Art. 44–47.

¹²⁶ *Ibid*, Art. 45.

¹²⁷ SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, p. 19.

Rights, provided that the States concerned have recognised the Court's contentious jurisdiction.¹²⁸ Only States Parties and the Inter-American Commission on Human Rights may submit cases to the Court, and proceedings before the Commission must first be completed. Where the Court finds a violation of Convention rights, it may order remedies, including the restoration of violated rights and the payment of compensation. In urgent situations, the Court may also adopt provisional measures to prevent irreparable harm.¹²⁹ In addition to its contentious jurisdiction, the Court exercises an important advisory function, allowing OAS Member States and certain OAS organs to request advisory opinions concerning the interpretation of human rights treaties and the compatibility of domestic laws with international human rights standards.¹³⁰ Judgments of the Court are final and binding, and States Parties undertake to comply with them.

From the perspective of AI-related human rights protection, this institutional structure is significant because it combines individual petition procedures, broader thematic monitoring and binding judicial remedies. The Commission may identify systemic or emerging patterns of technologically mediated harm, while the Court provides a judicial avenue for the authoritative interpretation and enforcement of Convention rights. At the same time, the indirect access of individuals to the Court and the dependence of the system on prior proceedings before the Commission may affect the speed with which effective judicial protection can be obtained in rapidly evolving technological contexts.

¹²⁸ *American Convention on Human Rights: "Pact of San José, Costa Rica"*, San José, 22 November 1969, Art. 62 (3).

¹²⁹ *Ibid*, Art. 63.

¹³⁰ *Ibid*, Art. 64.

5. CONTEMPORARY CHALLENGES AND THE IMPACT OF TECHNOLOGICAL CHANGE

Having outlined the conceptual and normative foundations, as well as institutional structure of IHRL, it is necessary to examine the extent to which this framework is capable of responding to contemporary technological developments. While IHRL provides an established, yet evolving normative framework, AI exposes and intensifies a number of pre-existing structural, interpretative and institutional limitations. The following sections therefore examine the principal challenges affecting the application and enforcement of existing human rights standards in AI-driven environments.

5.1. Structural Limitations of International Human Rights Law

Although IHRL has demonstrated a notable capacity to adapt over time, its normative design remains deeply rooted in a state-centred paradigm. This is reflected in the predominance of the territorial model for the application of human rights obligations, the primary focus on state responsibility for violations rather than on transnational corporations and other non-state actors, and the tendency of IHRL bodies to interpret and apply human rights instruments in relative isolations from other areas of law.¹³¹ These structural limitations are not entirely new, but rather stem from the historical foundations of the system, as well as the consent-based nature of international obligations, which constrains the enforcement of norms against non-consenting States. This structural configuration is further complicated by an inherent tension within the system, whereby States responsible for human rights violations remain directly involved in standard-setting, compliance monitoring, and enforcement processes.¹³² In the context of AI, these pre-existing limitations of IHRL are further pronounced. In the context of AI, these pre-existing limi-

¹³¹ SHANY, Y.: *Digital Rights and the Outer Limits of International Human Rights Law*. In: German Law Journal, Special Issue: International Law and Digitalization, Vol. 24, 2023, p. 464.

¹³² SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, pp. 28–29.

tations become particularly pronounced, as technological developments increasingly challenge traditional assumptions concerning territoriality, attribution and responsibility. AI therefore not only creates new human rights risks, but also exposes and intensifies structural weaknesses already embedded within the existing international human rights framework.

5.2. Opacity and Algorithmic Decision-Making

One of the most significant challenges posed by AI systems from the perspective of IHRL relates to their increasing opacity and the growing reliance on algorithmic decision-making processes in both the public and private sectors. A result, existing human rights norms are increasingly confronted with situations in which their application becomes uncertain or difficult.

One of the most notable characteristic of AI is its opacity, often described through the “black box” metaphor, meaning that the processes by which such systems transform input data into decisions may remain opaque even to their developers.¹³³ As *Almada* explains, this metaphor portrays AI systems as entities whose internal workings are not readily open to scrutiny, with such opacity arising from technical factors, including the vast quantities of data they process and their inherent complexity. It evokes a sense of risk and loss of control, as this opacity challenges the effective exercise of existing legal rights and safeguards.¹³⁴ At the same time, even where biased or discriminatory outcomes produced by AI systems can be identified, determining the precise origin of such bias often remains highly challenging.¹³⁵ It may be unclear whether the discriminatory result stems from the system’s own decision-making processes or whether it reflects biases embedded in the datasets, patterns of human behaviour, or societal inequalities on which the

¹³³ AWAD, E. et al.: *Crowd sourcing Moral Machines*. In: Communications of the ACM, Vol. 63, No. 3, 2020, p. 49. Online: <https://doi.org/10.1145/3339904>.

¹³⁴ ALMADA, M.: *Governing the Black Box of Artificial Intelligence*. In: *SSRN Electronic Journal*, January 2023, p. 2. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4587609.

¹³⁵ Council of Europe – Steering Committee for Human Rights: *CDDH Handbook on Human Rights and Artificial Intelligence*, CDDH(2025)R103 Addendum 3, 12 January 2026, para. 20. Online: <https://rm.coe.int/steering-committee-for-human-rights-cddh-cddh-handbook-on-human-rights/48802a2f59>.

system was trained¹³⁶ or with which it has interacted.¹³⁷ Algorithmic harms frequently operate at a systemic rather than an individual level: a biased system may affect thousands of individuals who share similar characteristics rather than producing a single identifiable violation.¹³⁸ Therefore, where data serve as the basis for decisions or evidentiary conclusions, the link between the processed information and the resulting outcome should remain sufficiently transparent to allow independent scrutiny, verification, and critical evaluation.¹³⁹

In many cases, meaningful explanations of AI-driven outcomes can only be inferred indirectly through analysing the relationship between the system's inputs and outputs, especially where developers or operators are unwilling or unable to disclose the internal functioning of the system.¹⁴⁰ Furthermore, AI systems are often protected by proprietary restrictions and characterised by considerable technical complexity, which may hinder even public authorities from fully understanding, explaining, or effectively auditing the systems they employ.

The significance of opacity therefore extends beyond the technical difficulty of understanding how an AI system operates. Where individuals cannot meaningfully understand the basis on which a decision affecting them was reached, their ability to challenge that decision, exercise procedural rights and obtain effective judicial or administrative review may be substantially weakened.

¹³⁶ GRAVETT, W.: *Sentenced by an algorithm — Bias and lack of accuracy in risk-assessment software in the United States criminal justice system*. In: *South African Journal of Criminal Justice*, Vol. 34, No. 1, p. 31.

¹³⁷ AWAD, E. et al.: *Crowd sourcing Moral Machines*. In: *Communications of the ACM*, Vol. 63, No. 3, 2020, p. 49.

¹³⁸ WACHTER, S., MITTELSTADT, B.D., RUSSELL, CH.: *Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI*. In: *Computer Law & Security Review*, Vol. 41, 105567, 2021, pp. 5–6.

¹³⁹ MITTELSTADT, B.D., ALLO, P., TADDEO, M., WACHTER, S., FLORIDI, L.: *The ethics of algorithms: Mapping the debate*. In: *Big Data & Society*, Vol. 3, No. 2, 2016, p. 4; UN General Assembly: *Promotion and protection of human rights in the context of digital technologies*, A/RES/78/213, 19 December 2023, para. 20 (c).

¹⁴⁰ CASTELLUCCIA, C., LE MÉTAYER, D.: *Understanding algorithmic decision-making: Opportunities and challenges*. European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA), European Parliament, March 2019, pp. 47–50.

5.3. Accountability and Governance Gap

Closely linked to the “black box” phenomenon is the dynamic and evolving nature of AI systems. Many AI systems continuously learn and adapt their decision-making processes over time, which may significantly complicate human efforts to understand, predict, and effectively regulate their potential negative impacts and externalities. More broadly, AI systems are often developed and deployed within short innovation cycles shaped by commercial incentives and rapid technological change, while the adoption and adaptation of adequate legal frameworks tends to proceed at a considerably slower pace.¹⁴¹ This temporal mismatch creates a persistent governance challenge, as legal safeguards may remain incomplete or outdated while AI systems are already deployed at scale. In the absence of timely and comprehensive regulation, human rights protection may therefore, at least temporarily, depend to a considerable extent on the internal governance practices and voluntary compliance of private actors. While such flexibility may facilitate rapid adaptation, it also raises concerns regarding consistency, accountability and the effectiveness of human rights protection in rapidly evolving technological contexts.¹⁴² Technological innovation is often driven primarily by technical and commercial considerations, while human rights standards are not always sufficiently integrated into the design and deployment of AI systems. As a consequence, decisions affecting the protection of human rights may increasingly be shaped by private actors and practical considerations rather than comprehensive legal assessment. This problem is further complicated by the proliferation of non-binding ethical principles, which frequently lack effective enforcement mechanisms. In connection with the above, it can be stated a further conceptual limitation arises from the uneven prioritisation of technological risks within current regulatory and scholarly discourse. Existing debates on the human rights implications of emerging technologies tend to focus predominantly on certain well-established areas, such as privacy, freedom of expression, disinformation, and online harms. While these issues are undoubtedly significant, this selective attention may result in the rel-

¹⁴¹ RAYFUSE, R.: *Public International Law and the Regulation of Emerging Technologies*. In: BROWNSWORD, R., SCOTFORD, E., YEUNG, K. (eds.): *The Oxford Handbook of Law, Regulation and Technology*, Oxford: Oxford University Press, 2017, pp. 500–501.

¹⁴² UN Human Rights Council: *Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/47/52, 19 May 2021, para. 52.

ative neglect of other equally important concerns. In particular, the broader structural impacts of digital technologies, including questions of disempowerment, social inequality, and the distribution of technological benefits and risks, remain comparatively underexplored.¹⁴³ This imbalance suggests that the current human rights discourse may not yet fully capture the systemic and long-term implications of AI and related technologies. A more comprehensive approach is therefore required, one that considers not only immediate interference with individual rights, but also structural questions concerning technological access, social inequality, participation and the distribution of decision-making power.

An important challenge associated with the use of AI concerns accountability, which is often dispersed among multiple actors, including developers, data providers, private companies, and public institutions. In practice, this fragmentation of responsibility may create situations in which it becomes difficult to identify who should be held legally or ethically accountable for harmful outcomes produced by AI systems. The complexity of AI supply chains, the involvement of multiple entities in the design, deployment and operation of AI systems, as well as the opacity of certain models, may result in accountability gaps in which responsibility is diluted or shifted between actors. These challenges become particularly significant in cases where AI systems affect fundamental rights, for example through discriminatory decision-making, unlawful surveillance, or erroneous automated decisions impacting individuals' access to employment, healthcare, social benefits, or justice. Against this background, effective AI governance requires clearer allocation of responsibilities, appropriate oversight mechanisms and legal frameworks capable of ensuring that accountability follows the distribution of actual decision-making power throughout the AI lifecycle.

5.4. Institutional and Operational Challenges

From the operational perspective, considerable challenge arises from the growing disparity between the increasing complexity of human rights challenges associated with emerging technologies and the limited resources available to the institutions tasked with addressing them. As digital and AI-driven systems generate more sophisticated and large-scale risks, human rights mechanisms are required to engage with technically demanding is-

¹⁴³ *Ibid*, para. 51.

sues that often exceed their existing capacities. This imbalance is particularly evident in the context of international human rights bodies, which are increasingly expected to address novel technological concerns while operating under constrained financial, technical, and human resources. As a result, these institutions may face difficulties in providing timely, consistent, and sufficiently specialised responses to technology-related human rights risks.¹⁴⁴ AI-related interferences are often systemic, technologically mediated and transnational, making them difficult to capture through complaint-based procedures and ex post adjudication, which remain the standard modes of operation for many existing human-rights institutions, including treaty bodies and regional courts.¹⁴⁵ These dynamics create what scholars describe as an “algorithmic accountability gap”: a structural mismatch between the ways in which harm emerges in AI-driven systems and the institutional frameworks traditionally designed to respond to human-rights violations.¹⁴⁶ Many human-rights mechanisms remain primarily reactive, whereas algorithmic harms often require preventive, continuous and technically informed oversight.¹⁴⁷ This situation raises broader concerns regarding the long-term sustainability and effectiveness of human rights protection in an increasingly complex technological environment.

Operational gaps are also closely connected to broader questions of international AI governance. While AI systems operate across borders and on a global scale, regulatory responses remain predominantly national or regional, resulting in a fragmented regulatory landscape marked by both overlaps and governance gaps.¹⁴⁸ The transnational nature of digital technologies further complicates the application of traditional jurisdictional principles, often limiting states’ ability to regulate harmful conduct or ensure effective rem-

¹⁴⁴ *Ibid*, para. 53.

¹⁴⁵ O’FLAHERTY, M.: *Human rights and the UN: practice before the treaty bodies*. Second Edition, Boston: Nijhoff: Brill, Law Specials, Vol. 54, 2021, 240 p.

¹⁴⁶ KROLL, J. et al.: Accountable Algorithms. In: *University of Pennsylvania Law Review*, Vol. 165, No. 3, 2017, pp 633–705; RAJI, I. D. & SMART, A.: Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. In: *FAT*20 Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, 2020, pp. 33–37; HORNEBER, D., LAUMER, S.: Algorithmic Accountability. In: *Business & Information Systems Engineering*, Vol. 65, No. 6, 2023, pp. 724–727.

¹⁴⁷ OHCHR: *Mapping report: human rights and new and emerging digital technologies Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/56/45, 20 August 2024, para. 61.

¹⁴⁸ MATTHIAS, A.: The responsibility gap: Ascribing responsibility for the actions of learning automata. In: *Ethics and Information Technology*, Vol. 6, 2004, pp. 181–182.

edies for individuals affected across multiple jurisdictions. These challenges highlight the growing need for stronger international coordination and co-operation in the governance of AI.¹⁴⁹ These challenges strengthen the case for enhanced international coordination and cooperation in AI governance.

These institutional limitations have also prompted renewed debate concerning the need for new oversight bodies or adapted institutional architectures capable of responding to the specific characteristics of AI systems. The relevant question is therefore not merely whether new institutions are required, but whether existing mechanisms can be equipped with the expertise, procedures and coordination tools necessary to perform their functions effectively in AI-related contexts.¹⁵⁰

5.5. Role of Private Actors

A further challenge concerns the growing prominence of the private sector in the development and deployment of AI. As a result, private actors increasingly exercise significant influence over the enjoyment of human rights. While important progress has been made in articulating corporate human rights responsibilities, particularly through frameworks such as the UN Guiding Principles on Business and Human Rights, significant gaps remain in their effective implementation. These limitations become especially significant in the AI context, where private companies may exercise substantial control over the design of systems, access to data, technical infrastructure and the conditions under which AI is deployed. Certain business models continue to exploit legal and regulatory grey areas, thereby circumventing or weakening existing safeguards. In this sense, technological innovation may not only expose existing gaps, but in some cases actively contribute to their expansion. These challenges are particularly evident in sectors such as digital labour platforms, where the absence of comprehensive regulatory frameworks raises concerns regarding the protection of workers' rights. More broadly, the increasing reliance on private actors in the governance of digital technologies highlights the need for stronger and more coherent mechanisms to ensure

¹⁴⁹ UN Human Rights Council: *Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/47/52, 19 May 2021, para. 54.

¹⁵⁰ See Part IV – *Institutional Frameworks for the Protection of Human Rights in the Age of Artificial Intelligence*.

that corporate practices are aligned with international human rights standards.¹⁵¹ Functions that were once traditionally associated with the State, such as identity verification and the management of personal data, are increasingly performed by private companies, including large digital platforms. As a result, these actors are able to collect and process vast amounts of personal information, often exceeding the data held by public authorities. This shift reflects a broader transformation in the governance of data, where private entities occupy a central position within so-called datafication processes.¹⁵² In this evolving landscape, the human rights responsibilities of private actors are becoming both more significant and more complex. This development further underscores the limitations of existing frameworks, which were not originally designed to address situations in which private entities exercise substantial influence over the conditions under which human rights are realised. The central challenge is therefore not simply to acknowledge the growing role of private actors, but to ensure that responsibility and accountability correspond to the degree of control and influence they exercise throughout the AI lifecycle.

¹⁵¹ UN Human Rights Council: *Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/47/52, 19 May 2021, para. 55.

¹⁵² *Ibid*, para. 56.

II. REGULATORY RESPONSES – FROM SOFT LAW TO BINDING FRAMEWORKS ON ARTIFICIAL INTELLIGENCE

The growing impact of AI systems on the enjoyment of human rights has prompted an increasingly complex regulatory response at both the global and regional levels.¹⁵³ As demonstrated in the preceding analysis, AI systems give rise not only to new forms of risk, but also to structural challenges that test the adequacy of existing legal frameworks. In response, a diverse range of regulatory approaches has emerged, reflecting varying institutional mandates, normative priorities, and levels of technological development. These responses encompass both non-binding instruments, which have played a pioneering role in articulating principles and standards for the governance of AI, and binding legal frameworks, which seek to translate these principles into enforceable obligations. Together, they illustrate the gradual formation of a complex regulatory landscape, characterised by both convergence and fragmentation.

Although existing norms of IHRL provide a certain foundational framework for the regulation of AI, a gradual scholarly and political consensus has emerged regarding the need for more specific legal responses reflecting the characteristics of AI as a complex technological phenomenon.¹⁵⁴ This evolving regulatory landscape is increasingly framed through a human rights lens. As emphasised by *Volker Türk*, any governance of AI must be firmly grounded in IHRL. This includes the need to assess human rights risks and impacts throughout the entire lifecycle of AI systems, to ensure transpar-

¹⁵³ Several States, particularly those with advanced technological capabilities, have also undertaken efforts to regulate artificial intelligence at the domestic level, adopting national legal acts or policy instruments governing the development and use of AI systems. Examples include countries such as Brazil, China, Japan, Saudi Arabia, Singapore, South Korea, Vietnam, as well as several States in the United States of America, such as California, Colorado, Utah. For more details see the OECD.AI Policy Navigator. Online: <https://oecd.ai/en/dashboards/policy-initiatives>.

¹⁵⁴ SHANY, Y.: *White Paper - The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025. Online: <https://afp.oxford-aiethics.ox.ac.uk/sitefiles/white-paper-professor-yuval-shany.pdf>; MITTELSTADT, B.: *Principles alone cannot guarantee ethical AI*. In: *Nature Machine Intelligence*, Vol. 1, 2019, pp. 501–507.

ency, independent oversight, and access to effective remedies, particularly where AI technologies are deployed by public authorities. At the same time, systems that cannot be operated in compliance with human rights standards may require restriction or even prohibition until adequate safeguards are in place. These considerations also extend beyond AI-specific regulation, highlighting the importance of effectively implementing existing legal frameworks, including data protection, competition law, and sector-specific regulations. Moreover, a human rights-based approach to AI governance requires not only state action, but also corporate responsibility, in line with established standards such as the UN Guiding Principles on Business and Human Rights.¹⁵⁵

This Part therefore examines the principal regulatory responses to AI, with particular attention to the extent to which they reflect and reinforce the protection of human rights and fundamental freedoms.

¹⁵⁵ OHCHR: *Artificial intelligence must be grounded in human rights, says High Commissioner*, 12 July 2023. Online: <https://www.ohchr.org/en/statements/2023/07/artificial-intelligence-must-be-grounded-human-rights-says-high-commissioner>.

6. SOFT LAW APPROACHES TO THE REGULATION OF ARTIFICIAL INTELLIGENCE

As is commonly the case in other areas of public international law, the initial regulatory responses to AI have taken the form of non-binding instruments. States, international intergovernmental organisations, and relevant stakeholders have approached the governance of AI primarily through the development of soft law frameworks, including principles, guidelines, and policy recommendations. This reliance on non-binding regulation can be attributed to several factors. On the one hand, it reflects the persistent conceptual uncertainty surrounding AI, including the absence of a universally accepted definition.¹⁵⁶ On the other hand, it is closely linked to the limited willingness of States and other actors to adopt binding legal obligations in a rapidly evolving technological field, where premature regulation may risk constraining innovation, creating regulatory asymmetries, or failing to accommodate future developments.

In parallel, international efforts have focused on the articulation of ethical principles intended to guide the development and use of AI systems.¹⁵⁷ In the absence of binding legal rules, such ethical frameworks have played a crucial role in identifying core normative standards. A general requirement for these frameworks is that they remain compatible with fundamental values such as human dignity, integrity, freedom, privacy, and cultural and gender diversity, as well as with the broader catalogue of internationally recognised human rights.¹⁵⁸ In this sense, ethical principles have functioned as a precursor to more formal legal regulation.

¹⁵⁶ In recent years, however, a gradual convergence in the definition of AI can be observed, particularly in international and regional instruments, which increasingly adopt broadly similar, function-oriented formulations of AI systems.

¹⁵⁷ Over 160 ethical guidelines on AI have been developed by international organizations, States, or relevant stakeholders. See KLUČKA, J.: *General Overview of the Artificial Intelligence and International Law*. In: KLUČKA, J., BAKOŠOVÁ, L., SISÁK, L. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Leges, 2021, p. 15.

¹⁵⁸ KLUČKA, J.: *Úloha a dôležitosť etických pravidiel v systémoch umelej inteligencie*. In: *Právny obzor*, Vol. 108, No. 3, 2025, p. 244. See also High-Level Expert Group on Artificial Intelligence: *Ethics Guidelines for Trustworthy AI*, Brussels, 2019, pp. 12–16. Online: <https://>

In this context, soft law has emerged as a flexible and adaptive regulatory tool, enabling the articulation of normative standards while preserving a degree of regulatory discretion. This chapter therefore examines the principal soft law approaches to AI governance, with particular attention to their contribution to the extent to which they incorporate human rights standards and contribute to the gradual development of a more coherent regulatory framework. Emphasis is placed on the initiatives developed within the framework of the UN, the OECD, as well as selected regional organisations that have progressively engaged in the regulation of AI.

6.1. The United Nations

At the universal level, the UN has emerged as a central actor in shaping the normative framework governing AI. Rather than adopting a single comprehensive legal instrument, the UN approach is characterised by the development of a diverse set of soft law standards, policy frameworks, and institutional initiatives. These instruments progressively connect AI governance with IHRL, sustainable development, international cooperation and broader questions of global digital governance.

Since 2019, a series of key UN reports and initiatives have contributed to shaping this emerging normative landscape. Early documents, such as *The Age of Digital Interdependence*¹⁵⁹ and the Secretary-General's *Roadmap for Digital Cooperation*, emphasised the need for enhanced global coordination, transparency, and accountability in the development and use of digital technologies, including AI.¹⁶⁰ Subsequent initiatives, including *Our Common Agenda*¹⁶¹ and the *Principles for the Ethical Use of Artificial Intelligence in the United Nations System*, further strengthened this trajectory by promoting human-centred and rights-based approaches grounded in fairness, trans-

digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai; UNESCO: *Recommendation on the Ethics of Artificial Intelligence*, SHS/BIO/PI/2021/1, 23 November 2021.

¹⁵⁹ UN Secretary-General's High-level Panel on Digital Cooperation: *The Age of Digital Interdependence*, 2019. Online: <https://www.un.org/en/pdfs/DigitalCooperation-reportfor%20web.pdf>.

¹⁶⁰ UN General Assembly: *Roadmap for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation: Report of the Secretary General*, A/RES/74/821, 29 May 2020.

¹⁶¹ United Nations: *Our Common Agenda: Report of the Secretary-General*, New York: United Nations, 2021, Online: https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Report_English.pdf.

parency, accountability, safety and human autonomy. At the same time, they highlight the need for safety, security, and oversight mechanisms to mitigate the risks associated with the deployment of AI systems.¹⁶²

In parallel to these policy-oriented initiatives, the normative framework has been further reinforced through resolutions of the UN General Assembly addressing the human rights implications of digital technologies, including AI. In particular, the resolution *Promotion and protection of human rights in the context of digital technologies* affirms that human rights must be respected, protected, and promoted throughout the entire life cycle of AI systems, while also recognising the potential of such technologies to support the realisation of human rights. Moreover, it calls upon States and other relevant actors to conduct human rights due diligence, including regular and comprehensive impact assessments, in order to prevent and mitigate adverse effects and to ensure effective remedies, human oversight, accountability, and legal responsibility. The resolution further emphasises the need to refrain from the use of AI systems that cannot be operated in compliance with IHRL or that pose disproportionate risks to the enjoyment of fundamental rights, unless adequate safeguards are in place. In addition, the resolution highlights the importance of inclusive and participatory governance, encouraging the involvement of all relevant stakeholders in decisions concerning the development and use of AI systems.¹⁶³

Building upon this framework, the UN General Assembly further developed its approach in the resolution *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*. Unlike earlier resolutions primarily focused on reaffirming the applicability of human rights in the digital environment, this instrument places greater emphasis on the governance and operationalisation of AI systems. It advances a more structured understanding of AI governance by promoting the development of regulatory frameworks, risk management mechanisms, and safeguards throughout the entire life cycle of AI systems. At the same time, it integrates human rights considerations into a broader policy context, linking AI governance with sustainable development objectives and the need to ad-

¹⁶² United Nations Inter-Agency Working Group on Artificial Intelligence: *Principles for the Ethical Use of Artificial Intelligence in the United Nations System*, 20 September 2022, pp. 2–6. Online: https://unsceb.org/sites/default/files/2022-09/Principles%20for%20the%20Ethical%20Use%20of%20AI%20in%20the%20UN%20System_1.pdf.

¹⁶³ UN General Assembly: *Promotion and protection of human rights in the context of digital technologies*, A/RES/78/213, 19 December 2023, para. 19–21.

dress global inequalities, including digital and technological divides.¹⁶⁴ The resolution thus reflects a shift from general normative affirmation towards more implementation-oriented guidance, highlighting the importance of international cooperation, capacity-building, and inclusive participation in shaping AI governance at the global level.

An important step towards a more structured approach to global AI governance was the establishment of the High-level Advisory Body on Artificial Intelligence by the UN Secretary-General in 2023. Composed of representatives from academia, industry, governments, and civil society, the body was tasked with developing recommendations aimed at ensuring that AI governance promotes the common good while remaining aligned with human rights and the Sustainable Development Goals.¹⁶⁵ Its final report, *Governing AI for Humanity* (2024), identifies key structural challenges in the current governance landscape, including regulatory fragmentation, unequal access to AI technologies, and risks related to bias and accountability.¹⁶⁶ Moreover, it advances proposals for strengthening global coordination, including the creation of new institutional mechanisms and frameworks for data governance, capacity-building, and international cooperation.¹⁶⁷ These recommendations subsequently informed broader UN initiatives, including the adoption of the *Global Digital Compact*.

The Global Digital Compact further reinforces the centrality of human rights within the emerging framework of AI governance. It explicitly affirms that the development and deployment of digital technologies, including AI, must remain aligned with international law, in particular with IHRL.¹⁶⁸ At the same time, the Compact adopts a broader perspective by linking human rights considerations with issues of digital inclusion, equitable access, and sustainable development. It highlights the importance of bridging digital divides, promoting inclusive participation in decision-making processes, and

¹⁶⁴ UN General Assembly: *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*, A/RES/78/265, 21 March 2024, para. 3–6.

¹⁶⁵ United Nations: *UN Secretary-General launches AI Advisory Body on risks, opportunities, and international governance of artificial intelligence*, Press release, 2023. Online: https://www.un.org/sites/un2.un.org/files/231025_press-release-aiab.pdf.

¹⁶⁶ United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, pp. 42–46.

¹⁶⁷ *Ibid.*, pp. 47–76.

¹⁶⁸ UN General Assembly: *Pact for the Future – Annex – Global Digital Compact*, A/RES/79/1, 22 September 2024, para. 8 (c) and 22.

ensuring that the benefits of AI are distributed in a fair and non-discriminatory manner.

This human rights-oriented approach is further elaborated in Objective 3 of the Compact, which calls for the creation of an inclusive, open, and safe digital environment. In this context, it emphasises the obligation to respect, protect, and promote human rights throughout the entire life cycle of digital technologies, while also recognising the responsibilities of non-State actors, in particular private companies, in line with the UN Guiding Principles on Business and Human Rights. Particular importance is attached to the implementation of human rights due diligence, the establishment of effective oversight and remedy mechanisms, and the need to ensure that national regulatory frameworks governing digital technologies remain consistent with international human rights obligations.¹⁶⁹ Rather than introducing new rights, the Compact reinforces the applicability of existing human rights standards in the digital environment, while adapting them to the specific risks and opportunities associated with AI.

Taken together, these initiatives illustrate the gradual emergence of a complex yet predominantly non-binding framework for AI governance within the UN system. While this approach allows for flexibility and rapid normative development, it also raises important questions regarding coherence, enforceability, and the overall effectiveness of international regulation. At the same time, the Compact has been subject to criticism, particularly with regard to its human rights and gender narratives, which are sometimes seen as reflecting predominantly Western perspectives and thus potentially insufficiently accommodating diverse cultural, social, and regional contexts.¹⁷⁰ The principal challenge for the UN framework therefore lies not in the absence of normative guidance, but in translating an increasingly dense body of soft-law standards into coherent and practically effective governance arrangements.

¹⁶⁹ *Ibid*, para. 22–25.

¹⁷⁰ See for instance KOMAITIS, K.: *Analysis: a brave new reality after the UN's Global Digital Compact*, The Digital Forensic Research Lab, 1 October 2024. Online: <https://dfrlab.org/2024/10/01/analysis-a-brave-new-reality-after-the-uns-global-digital-compact/>; HENDRIX, J.: *Reactions to the Adoption of the UN Global Digital Compact*, TechPolicy. Press, 8 October 2024. Online: <https://www.techpolicy.press/reactions-to-the-adoption-of-the-un-global-digital-compact/>.

6.1.1. Human Rights-Based Approaches to AI Governance within the UN System

Over the past two decades, the UN Human Rights Council (hereinafter “HRC”) has become an important forum for addressing the human rights implications of emerging technologies through resolutions, thematic reports, and the work of its subsidiary mechanisms, particularly the Special Procedures.¹⁷¹ In recent years, its attention has increasingly focused on digital technologies and AI, especially regarding issues such as privacy, surveillance, disinformation, and the broader societal impacts of data-driven systems,¹⁷² thereby contributing to the gradual emergence of a human rights-based approach to AI governance.

Early normative attention within the UN human rights framework focused primarily on the protection of the right to privacy in the digital age,¹⁷³ particularly through recurring resolutions affirming that the same rights individuals enjoy offline must also be protected online.¹⁷⁴ This principle subsequently became one of the normative foundations for later discussions concerning AI systems and data-driven technologies. Over time, however, the Council expanded its focus to encompass broader issues connected with digital governance, including freedom of opinion and expression, access to information, algorithmic discrimination, biometric surveillance, and auto-

¹⁷¹ UN Human Rights Council: *Non-take-up of rights in the context of social protection: Report of the Special Rapporteur on extreme poverty and human rights*, Olivier De Schutter, A/HRC/RES/50/38, 19 April 2022; UN Human Rights Council: *Rights of persons with disabilities: Report of the Special Rapporteur on the rights of persons with disabilities*, A/HRC/RES/49/52, 28 December 2021.

¹⁷² UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/51/17, 4 August 2022; UN Human Rights Council: *Impact of new technologies on the promotion and protection of human rights in the context of assemblies, including peaceful protests* Report of the United Nations High Commissioner for Human Rights, A/HRC/RES/44/24, 24 June 2020; UN Human Rights Council: *Impact of disinformation on the enjoyment and realization of human rights: Report of the Human Rights Council Advisory Committee*, A/HRC/RES/61/68, 4 February 2026.

¹⁷³ UN Human Rights Council: *Right to privacy in the digital age*, A/HRC/RES/48/4, 7 October 2021; UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/RES/60/45, 29 August 2025.

¹⁷⁴ E.g. UN Human Rights Council: *New and emerging digital technologies and human rights*, A/HRC/RES/47/23, 13 July 2021, Preamble; UN Human Rights Council: *Right to privacy in the digital age*, A/HRC/RES/54/21, 12 October 2023, para. 4.

mated decision-making. Particular attention has been devoted to the risks associated with large-scale data processing and the increasing use of AI-driven technologies by both public authorities and private actors.¹⁷⁵

The Council's Special Procedures have further contributed to the normative development of this field. Mandate holders, including the Special Rapporteurs on the right to privacy,¹⁷⁶ on the rights of persons with disabilities,¹⁷⁷ and contemporary forms of racism,¹⁷⁸ have examined the implications of AI and digital technologies in their thematic reports, thereby providing more detailed guidance concerning surveillance technologies, content moderation, automated decision-making, biometric technologies, and corporate accountability in the digital environment. Their thematic flexibility is particularly valuable in the context of AI, as it enables emerging technological risks to be addressed within existing human rights mandates without requiring the prior creation of AI-specific institutional mechanisms.

Within this broader normative ecosystem, the OHCHR has assumed an increasingly important role as both a coordinating and interpretative actor. Through thematic reports, policy guidance, and practical implementation initiatives, the OHCHR has contributed to translating existing human rights standards into the context of emerging digital technologies and AI systems. Its reports on the right to privacy in the digital age have been particularly influential in articulating concerns relating to profiling, automated decision-making, facial recognition technologies, and the broader implications of AI-driven surveillance.¹⁷⁹ The OHCHR has further emphasised the need for comprehensive and non-fragmented regulatory frameworks addressing

¹⁷⁵ UN General Assembly: *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, A/RES/73/348, 29 August 2018; UN Human Rights Council: *New and emerging digital technologies and human rights*, A/HRC/RES/47/23, 13 July 2021.

¹⁷⁶ UN Human Rights Council: *Report of the Special Rapporteur on the right to privacy*, Joseph A. Cannataci, A/HRC/RES/34/60, 24 February 2018.

¹⁷⁷ UN Human Rights Council: *Rights of persons with disabilities: Report of the Special Rapporteur on the rights of persons with disabilities*, A/HRC/RES/49/52, 28 December 2021.

¹⁷⁸ UN Human Rights Council: *Racial and xenophobic discrimination and the use of digital technologies in border and immigration enforcement: Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, E. Tendayi Achiume, A/HRC/RES/48/76, 17 December 2021; UN Human Rights Council: *Racial discrimination and emerging digital technologies: a human rights analysis: Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, A/HRC/RES/44/57, 18 June 2020.

¹⁷⁹ UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/51/17, 4 August 2022;

discrimination and inequality in data collection and processing, while cautioning against broad exemptions for law enforcement, migration management, and national security that may undermine the effective protection of human rights. It has also advocated continuous human rights due diligence throughout the lifecycle of AI systems, alongside greater transparency, independent oversight, and safeguards against bias and discriminatory impacts in automated decision-making.¹⁸⁰ The growing role of the OHCHR has also been recognised in recent UN policy initiatives concerning digital governance. Both the Pact for the Future and the Global Digital Compact emphasise the need for closer coordination among UN actors working on digital technologies and human rights, while simultaneously reaffirming the OHCHR's role in providing expert guidance and normative support concerning the human rights implications of emerging technologies. At the same time, these instruments implicitly acknowledge the increasing institutional complexity and fragmentation characterising the UN digital governance landscape.¹⁸¹

Beyond its interpretative role, the OHCHR has also sought to operationalise human rights standards in the digital sphere through a range of practical initiatives and guidance instruments. A notable example is the B-Tech Project¹⁸², which seeks to address human rights risks arising from the activities of technology companies. The project seeks to translate existing human rights standards into practical guidance for States, businesses, and other stakeholders, thereby strengthening the implementation of the “protect, respect and remedy” framework in the digital environment.¹⁸³ Through the development of policy recommendations, stakeholder consultations, and thematic outputs, B-Tech promotes greater corporate accountability and encourages the integration of human rights due diligence into the design, development, and use of digital technologies. At the same time, the B-Tech Project highlights systemic challenges, including regulatory gaps, power asymmetries between

UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/RES/60/45, 29 August 2025.

¹⁸⁰ UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/RES/60/45, 29 August 2025, para. 53 (a), (e), 54 (b).

¹⁸¹ UN General Assembly: *Pact for the Future*, A/RES/79/1, 22 September 2024, para. 74 (b).

¹⁸² Business and Human Rights in Technology Project.

¹⁸³ OHCHR: *Business and Human Rights in Technology Project (“B-Tech Project”): Applying the UN Guiding Principles on Business and Human Rights to Digital Technologies*, 25 January 2020, p. 1. Online: <https://www.ohchr.org/sites/default/files/Documents/Issues/Business/B-Tech/BTechprojectoverview.pdf>.

States and private actors, and the transnational nature of digital platforms, which complicate effective oversight and enforcement.¹⁸⁴

Similarly, the OHCHR has developed the *Guidance of the Secretary-General: Human Rights Due Diligence for Digital Technology Use*, whose aim is to support all UN entities in implementing and strengthening human rights due diligence policies,¹⁸⁵ processes and practices in the use of digital technologies, including AI. Beyond its active engagement with the private sector, the OHCHR has also been instrumental in developing a range of guidance documents addressed to UN entities, focusing on digital technologies more broadly, while remaining directly relevant to the governance of AI. These efforts reflect the progressive embedding of human rights due diligence within international standards, as well as its increasing incorporation into national and supranational policy and regulatory frameworks.¹⁸⁶ Within this framework, human rights due diligence is conceptualised as a continuous and risk-based process requiring actors to identify, assess, prevent, mitigate, and address adverse human rights impacts throughout the lifecycle of digital technologies. Emphasis is placed on transparency, stakeholder participation, inclusion, intersectionality, and context-sensitive approaches capable of responding to the evolving societal impacts of AI systems. At the same time, these initiatives stress that human rights due diligence should not function merely as a formal compliance exercise, but rather as an ongoing governance mechanism aimed at achieving positive human rights outcomes for affected individuals and communities.¹⁸⁷

The activities of the HRC and the OHCHR collectively demonstrate that existing human rights institutions possess a considerable capacity to adapt established norms to technological change through interpretation, thematic guidance and operational tools. This adaptability represents an important strength of the existing IHRL framework. At the same time, however, this framework remains characterised by a degree of fragmentation,¹⁸⁸ consisting primarily of thematic soft-law instruments and sector-specific initiatives

¹⁸⁴ *Ibid.*, p. 2.

¹⁸⁵ Human Rights Due Diligence is a process that enables an entity to identify and address its human rights risks and impacts effectively.

¹⁸⁶ United Nations: *Guidance of the Secretary-General: Human Rights Due Diligence for Digital Technology Use: Guidance of the Secretary-General*, 1 May 2024, p. 2. Online: digital-technology-use-guidance-sg-1-en.pdf.

¹⁸⁷ *Ibid.*, p. 8.

¹⁸⁸ BALCERZAK, M.: *Implications of the United Nations human rights standards for the development of artificial intelligence*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J.

rather than a single coherent regulatory model. While this flexible approach enables responsiveness to rapid technological developments, it simultaneously raises questions concerning the consistency, effectiveness, and enforceability of existing international human rights responses to the systemic challenges posed by AI technologies.

In addition to the activities of the HRC and the OHCHR, important interpretative guidance relevant to AI governance has also emerged from the practice of UN treaty bodies, particularly through their general comments. Although formally non-binding, these instruments play an important role in clarifying the scope of treaty obligations and increasingly address issues related to digital technologies, privacy, automated decision-making, discrimination, and children's rights in the digital environment. Particularly relevant in this regard are *General Comment No. 25 of the Committee on the Rights of the Child on children's rights in relation to the digital environment*, *General Comment No. 36 of the Human Rights Committee on the right to life*, and *General Comment No. 24 of the Committee on Economic, Social and Cultural Rights concerning State obligations in the context of business activities*. Their practice further illustrates how existing treaty norms may be progressively interpreted in response to technological developments without requiring formal amendment of the underlying human rights instruments.

6.1.2. The United Nations Educational, Scientific and Cultural Organization

In contrast to the broader and more fragmented approach of the UN system, UNESCO has adopted a more structured normative framework through its *Recommendation on the Ethics of Artificial Intelligence*. This instrument represents the first global standard-setting initiative specifically dedicated to the ethical governance of AI. It approaches AI ethics as a systematic normative reflection, based on a holistic, comprehensive, multicultural and evolving framework of interdependent values, principles and actions that can guide societies in dealing responsibly with the known and unknown impacts of AI technologies on human beings, societies and the environment and ecosystems, and offers them a basis to accept or reject AI technologies. It considers ethics as a dynamic basis for the normative evaluation and guidance of AI

(eds): Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World. Cheltenham: Edward Elgar Publishing Limited, 2024, p. 8.

technologies, referring to human dignity, well-being and the prevention of harm as a compass and as rooted in the ethics of science and technology.¹⁸⁹

The Recommendation aims to establish a universal framework of values, principles, and actions to guide both States and non-State actors in the development and use of AI systems. It seeks to ensure that ethical considerations are embedded throughout the entire life cycle of AI, with particular emphasis on the protection of human rights, human dignity, and equality, including gender equality. At the same time, it promotes the safeguarding of the environment, biodiversity, and the interests of present and future generations, as well as the respect for cultural diversity. In addition, the Recommendation encourages inclusive, multi-stakeholder dialogue on ethical issues related to AI and seeks to promote equitable access to AI technologies and their benefits, with particular attention to developing countries.¹⁹⁰

The Recommendation is grounded in a set of core values, including the respect for, protection and promotion of human rights and fundamental freedoms, as well as human dignity. It further emphasises the importance of environmental and ecosystem sustainability, the promotion of diversity and inclusiveness, and the aspiration to foster peaceful, just, and interconnected societies. At the core of this framework lies the concept of human dignity, understood as the foundation of the universal and interdependent system of human rights and fundamental freedoms. Human rights are to be respected, protected, and promoted throughout the entire life cycle of AI systems, in accordance with international law, including IHRL. The framework further emphasises that AI systems must not harm, subordinate, or objectify individuals or communities in any dimension of their existence. Instead, their deployment should contribute to the enhancement of human well-being, while preserving individual autonomy and preventing any infringement of fundamental rights. Particular attention is devoted to vulnerable groups, whose interaction with AI systems requires heightened safeguards.¹⁹¹ In this sense, technological development is framed as a means of strengthening, rather than undermining, the protection and effective exercise of human rights.

The Recommendation is explicitly grounded in human rights principles, emphasising do no harm principle, fairness, non-discrimination, transparency, and accountability, while also incorporating broader considerations

¹⁸⁹ UNESCO: *Recommendation on the Ethics of Artificial Intelligence*, SHS/BIO/PI/2021/1, 23 November 2021, para. 1.

¹⁹⁰ *Ibid*, para. 8.

¹⁹¹ *Ibid*, para. 13–16.

such as environmental sustainability and social well-being.¹⁹² Unlike many other soft law instruments, it provides relatively detailed guidance on the implementation of these principles, including through ethical impact assessments, governance mechanisms, and policy measures at the national level.¹⁹³

Building upon the abovementioned Recommendation, UNESCO published the *Readiness Assessment Methodology* in 2023, which serves as a comprehensive evaluative framework for assessing the preparedness of individual states to govern AI in an ethical and human rights-compliant manner. The methodology examines a broad range of factors, including regulatory structures, institutional capacities, technological infrastructure, education, research, and broader socio-economic conditions, thereby enabling states to identify both regulatory deficiencies and areas requiring further capacity-building.¹⁹⁴

In parallel, UNESCO introduced the *Ethical Impact Assessment* tool, designed primarily to support public authorities and private actors involved in the development, procurement, deployment, and use of AI systems. Rather than formulating abstract ethical guidance, the tool translates the principles contained in the Recommendation into a structured decision-making framework based on practical questions and evaluative criteria relating to issues such as safety, transparency, fairness, non-discrimination, diversity, accountability, and human oversight.¹⁹⁵ In this way, it seeks to embed ethical and human rights considerations directly into the lifecycle of AI systems and into institutional decision-making processes concerning their deployment and use.

Taken together, these instruments do not seek to establish a separate or competing model of AI governance. Instead, they reinforce the broader human rights-oriented approach underpinning the UNESCO Recommendation by treating human rights as a foundational component of ethical AI

¹⁹² *Ibid*, para. 25–47; BALCERZAK, M.: *Implications of the United Nations human rights standards for the development of artificial intelligence*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, pp. 21–22.

¹⁹³ UNESCO: *Recommendation on the Ethics of Artificial Intelligence*, SHS/BIO/PI/2021/1, 23 November 2021, para. 131.

¹⁹⁴ UNESCO: *Readiness Assessment Methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence*, SHS/REI/BIO/REC-AIETHICS-TOOL/2023 Rev., 2023, pp. 5–6.

¹⁹⁵ UNESCO: *Ethical Impact Assessment: a tool of the Recommendation on the Ethics of Artificial Intelligence*, SHS/REI/BIO/REC-AIETHICS-TOOL-EIA/2023, 2023, pp. 18–42.

governance. Through these initiatives, UNESCO contributes to the gradual development of practical governance mechanisms intended to strengthen the capacity of states and institutions to address the societal and human rights challenges associated with emerging AI technologies.¹⁹⁶

6.1.3. The United Nations Children's Fund

Children constitute one of the most vulnerable groups of people in the context of the deployment and use of AI systems, as their evolving capacities and limited ability to understand or contest automated processes expose them to distinct risks associated with algorithmic decision-making. This vulnerability is further exacerbated by the fact that existing AI governance frameworks are predominantly designed for adult users and therefore insufficiently account for the specific rights and needs of children.¹⁹⁷

In response to the increasing impact of AI on children's lives, the United Nations Children's Fund (hereinafter "UNICEF"), as a key institution dedicated to the protection and promotion of children's rights,¹⁹⁸ has developed *Guidance on AI and children*. Grounded in the *Convention on the Rights of the Child*, the guidance advances a child-centred approach to AI governance, according to which AI policies and systems should protect children, ensure the equitable fulfilment of their rights and needs, and support their meaningful participation in an AI-driven world.¹⁹⁹ Its development reflects a progres-

¹⁹⁶ BALCERZAK, M.: *Implications of the United Nations human rights standards for the development of artificial intelligence*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, p. 23.

¹⁹⁷ WANG, G., ZHAO, J., VAN KLEEK, M., SHADBOLT, N.: *Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children*. Association for Computing Machinery CHI'22 Conference, April 29-May 5, 2022, New Orleans, LA, USA, p. 5. Online: <https://dl.acm.org/doi/pdf/10.1145/3491102.3502057>; PENAGOS, M., KASSIR, S., VOSLOO, S.: *Policy Brief on National AI Strategies and Children: Reviewing the Landscape and Identifying Windows of Opportunity*, Office of Global Insight and Policy, 2020, pp. 2–3. Online: <https://www.unicef.org/innocenti/media/2516/file/UNICEF-Global-Insight-national-AI-strategy-review-policy-brief.pdf>.

¹⁹⁸ The transformation of UNICEF's mandate from emergency assistance to a comprehensive rights-based approach has positioned it as a key actor in advancing the well-being and development of children within the international legal and policy framework.

¹⁹⁹ UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, p. 2. Online: <https://www.unicef.org/innocenti/media/11991/file/UNICEF-Innocenti-Guidance-on-AI-and-Children-3-2025.pdf>.

sively iterative and inclusive process combining normative reflection with empirical validation.

The initial draft (version 1.0, 2020) was subjected to extensive multi-stakeholder consultation, involving governments, international organizations, civil society, private-sector actors, and end-users, including children themselves. At this stage, the guidance primarily fulfilled an agenda-setting function, aiming to stimulate discussion and identify key normative gaps at the intersection of AI and children's rights.²⁰⁰ The feedback obtained was systematically incorporated into version 2.0 (2021), which marked a shift from conceptual framing towards a more structured and implementation-oriented framework. This phase was further reinforced by pilot initiatives involving governmental and business actors, enabling the testing of the guidance in concrete regulatory and operational settings. The insights derived from these pilots informed both subsequent revisions and the development of targeted practical tools designed to facilitate its application.²⁰¹

The most recent iteration (version 3, 2025) reflects a shift towards a more detailed and operational framework aligned with evolving international child rights standards, including General Comment No. 25 (2021) of the Committee on the Rights of the Child,²⁰² relevant UN General Assembly resolutions,²⁰³ and the Global Digital Compact. Its development was driven by rapid advances in AI technologies, particularly generative AI, and the growing integration of AI into children's everyday lives. The updated version addresses a broader range of risks and opportunities, including AI companions, harmful datasets, environmental impacts, AI-generated child sexual abuse material, and the use of AI in armed conflict, while also recognising AI's potential benefits for education, accessibility, and child well-being.²⁰⁴ Collectively, these developments illustrate the gradual evolution of the in-

²⁰⁰ UNICEF: *Policy guidance on AI for children: Draft 1.0*, New York: Office of Global Insight and Policy, 2020. Online: <https://www.unicef.org/innocenti/media/1326/file/UNICEF-Global-Insight-policy-guidance-AI-children-draft-1.0-2020.pdf>.

²⁰¹ UNICEF: *Policy guidance on AI for children: Draft 2.0*, New York: Office of Global Insight and Policy, 2021. Online: <https://www.unicef.org/innocenti/media/1341/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf>.

²⁰² Committee on the Rights of the Child: *General Comment No. 25 (2021) on children's rights in relation to the digital environment*, (CRC/C/GC/25), 2 March 2021.

²⁰³ UN General Assembly: *Promotion and protection of the rights of children*, A/C.3/78/L.19/Rev.1, 8 November 2023.

²⁰⁴ UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, p. 5.

strument from a primarily consultative initiative into a more structured and quasi-regulatory framework within international human rights governance.

Against this background, the guidance establishes ten core requirements for child-centred AI, structured around the principles of protection, provision, and participation. These principles are intended to guide all actors involved in the design, development, and deployment of AI systems. The principle of protection emphasises safeguarding children from harm, discrimination, and exploitation arising from AI systems. Provision focuses on ensuring that AI supports children's rights and well-being, particularly in areas such as education and health, while participation requires that children be meaningfully involved in decisions affecting their use of and interaction with AI technologies.²⁰⁵

Considering the above, UNICEF calls upon States, competent authorities and business actors to comply with a set of core requirements. First, it emphasises the need to establish appropriate regulatory frameworks, oversight and compliance mechanisms for child-centred AI. States are encouraged to review and update existing legislation to ensure full integration of children's rights, including through dedicated or existing oversight bodies and accessible mechanisms for redress.²⁰⁶ At the same time, AI systems deployed by governments must operate in conformity with both AI-specific regulation and existing child rights law.²⁰⁷

The second requirement concerns ensuring children's safety in the context of AI systems. It calls for the use of child rights impact assessments across all stages of the AI lifecycle, as well as continuous monitoring and evaluation.²⁰⁸ AI systems should be rigorously tested to ensure safety, security and robustness, with particular attention to harms arising from AI-generated or amplified content, such as misinformation, cyberbullying and online exploitation.²⁰⁹ It further highlights the need to prevent AI-enabled crimes against

²⁰⁵ *Ibid*, p. 8.

²⁰⁶ See *Convention on the Rights of the Child*, New York, 20 November 1989, UNTS Vol. 1577, p. 3.

²⁰⁷ UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, pp. 12–14.

²⁰⁸ Committee on the Rights of the Child: *General Comment No. 25 (2021) on children's rights in relation to the digital environment*, (CRC/C/GC/25), 2 March 2021, para. 23 and 38.

²⁰⁹ See for instance ITU: *Child Online Protection Guidelines for children, parents and educators, industry and policy-makers*, 2020. Online: <https://www.itu-cop-guidelines.com/>.

children and to address emerging risks related to AI-driven applications, including chatbots and virtual companions.²¹⁰

The requirement to protect children's data and privacy emphasises responsible data governance, including the adoption of a privacy-by-design approach. Particular attention should be paid to data minimisation and retention, ensuring that children's data is not used in ways that may negatively affect them later in life. Protection should also extend to group-based risks, while promoting children's data agency, including their ability to access, understand and control their personal data²¹¹ in line with their evolving capacities.²¹²

Ensuring non-discrimination and fairness requires addressing bias throughout the AI lifecycle, from data collection to deployment. Stakeholders must take proactive measures to prevent discriminatory outcomes and to support children in vulnerable situations, ensuring that AI systems contribute to reducing, rather than exacerbating, existing inequalities and promote equitable access to their benefits.²¹³

The requirement to ensure transparency, explainability and accountability calls for AI systems to be designed in a manner that is intelligible and accessible to children. Systems must clearly disclose their non-human nature in an age-appropriate way and provide explanations adapted to children's level of understanding. Responsibility for ensuring safe and rights-compliant AI cannot be shifted onto children, and clear accountability mechanisms must be established to monitor compliance and provide remedies.²¹⁴

Respect for children's rights through responsible AI practices requires a rights-based approach across the entire AI value chain.²¹⁵ In this regard, Gen-

²¹⁰ UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, pp. 14–22.

²¹¹ MONTGOMERY, K. C., CHESTER, J., KOPP, K.: *Data governance for young people in the commercialized digital environment*, Issue Brief, UNICEF, 2020. Online: UNICEF-Global-Insight-data-governance-commercialization-issue-brief-2020.pdf.

²¹² UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, pp. 22–25.

²¹³ *Ibid*, pp. 25–26.

²¹⁴ *Ibid*, pp. 27–28.

²¹⁵ OHCHR, UNICEF: *Taking a Child Rights-Based Approach to Implementing the UNGPs in the Digital Environment: A contribution to the B-Tech Project*, November 2024. Online: <https://www.unicef.org/childrightsandbusiness/media/1406/file/B-Tech-UNICEF-Briefing.pdf>.

eral Comment No. 16 (2013) on State obligations regarding the impact of the business sector on children's rights is of particular relevance. It frames States' obligations in relation to business activities affecting children's rights through the duties to respect, protect, and fulfil, complemented by the obligation to ensure effective remedies. The duty to respect requires States to refrain from interfering with children's rights and to avoid supporting harmful business practices. The duty to protect entails adopting and enforcing measures to prevent, investigate, and redress abuses by business actors. The duty to fulfil requires positive action to create an enabling environment for the realization of children's rights. Finally, States must ensure access to effective, child-sensitive remedies that provide timely and adequate reparation.²¹⁶ The Guidance further draws attention to specific risks associated with digital technologies, including child labour and exploitation, and emphasises the need for adequate capacity-building among all actors involved in the development and deployment of AI systems. It also highlights the potential to align market incentives with the promotion of trustworthy and rights-compliant AI.²¹⁷

The requirement to support children's best interests, development and well-being calls for AI systems to prioritise the holistic development of the child. This includes embedding a child rights-by-design approach, developing appropriate indicators of well-being,²¹⁸ and addressing environmental impacts of AI systems, recognising their indirect effects on children's rights. AI should also be leveraged to support sustainability and responsible development.²¹⁹

Ensuring inclusion of and for children requires meaningful participation of children in AI governance and system design, in line with their evolving capacities. It also calls for inclusive design approaches that reflect the diversity of children's needs and circumstances, as well as diversity among actors

²¹⁶ Committee on the Rights of the Child: *General Comment No. 16 (2013) on State obligations regarding the impact of the business sector on children's rights*, CRC/C/GC/16, 17 April 2013, para. 26–31.

²¹⁷ UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, pp. 28–31.

²¹⁸ UNICEF: *Children's Best Interests in Digital Policy and Practice: From the voices of children, youth and experts around the world*, April 2026, p. 70. Online: <https://www.unicef.org/innocenti/media/12841/file/UNICEF-Innocenti-Children's-Best-Interests-Digital-Policy-Practice-report-2026.pdf>.

²¹⁹ UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, pp. 32–35.

involved in the AI ecosystem.²²⁰ Particular attention must also be paid to the participatory dimension of children's rights, as articulated in *General Comment No. 12 (2009) on the right of the child to be heard*. This General Comment affirms that children have the right to express their views freely in all matters affecting them and to have those views given due weight in accordance with their age and maturity. It further emphasises that this right constitutes a fundamental principle of *the Convention on the Rights of the Child*, requiring States not only to ensure opportunities for expression but also to establish processes that enable meaningful and effective participation.²²¹ In the context of AI and digital technologies, this entails the inclusion of children's perspectives in decision-making processes related to the design, deployment, and governance of such systems. However, despite its normative clarity, the effective operationalisation of this right in AI governance remains limited, with children's participation often overlooked or reduced to formalistic consultation rather than substantive engagement.

Preparing children for present and future AI developments requires strengthening AI literacy and relevant skills through formal and informal education. This includes supporting educators, providing guidance on the responsible use of AI in education, and promoting broader societal awareness among parents and caregivers. At the same time, attention must be given to the impact of AI on future labour markets and children's long-term opportunities.²²²

Finally, creating an enabling environment for child-centred AI requires investment in digital infrastructure, efforts to bridge the digital divide, and support for research on AI for and with children. It also calls for enhanced international cooperation, multi-stakeholder governance, and the development of interoperable frameworks to ensure consistent protection and opportunities for children across jurisdictions.²²³

In addition to the above requirements, the guidance identifies a set of considerations that should inform AI policymaking and development activities. First, the implementation of these principles must be adapted to national and local contexts, considering varying levels of technological development,

²²⁰ *Ibid*, pp. 35–37.

²²¹ Committee on the Rights of the Child: *General Comment No. 12 (2009) on the right of the child to be heard*, CRC/C/GC/12, 20 July 2009, para. 40–45.

²²² UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, pp. 37–42.

²²³ *Ibid*, pp. 42–45.

regulatory maturity and the specific conditions under which AI systems are deployed. At the same time, such contextualisation must not undermine the universal protection of children's rights but rather ensure their effective realisation across diverse settings. Furthermore, the guidance emphasises the importance of foresight in AI governance, calling for forward-looking, anticipatory and adaptive approaches capable of responding to the rapidly evolving technological landscape. In particular, the increasing integration of AI into everyday environments, including through interconnected systems and emerging technologies such as neurotechnologies, requires governance frameworks that extend beyond traditional, screen-based interactions.

In this regard, States are encouraged to establish institutional mechanisms, such as horizon-scanning bodies or specialised ethics committees, tasked with monitoring technological developments and supporting proactive regulatory responses. Finally, the guidance underscores the need to ensure that legal and policy frameworks are future-ready. Rather than reacting retrospectively to technological change, governance approaches should be designed to remain applicable in the face of evolving technologies, including by focusing on the impacts of AI systems rather than their specific technical characteristics. Such an approach enhances the resilience and longevity of regulatory frameworks, while ensuring that the best interests of the child remain central in increasingly complex socio-technical environments.²²⁴

In sum, the UNICEF Guidance demonstrates how existing international human rights standards can be progressively adapted to the specific risks and needs arising in AI-driven environments without requiring the creation of an entirely separate normative framework. Its particular contribution lies in translating established children's rights into concrete governance requirements applicable throughout the AI lifecycle. At the same time, its effectiveness ultimately depends on the extent to which these non-binding standards are incorporated into domestic regulation, institutional practice and the conduct of private actors.

6.2. The Organization for Economic Co-operation and Development

Alongside the UN, the Organization for Economic Co-operation and Development (OECD) has played a particularly influential role in the develop-

²²⁴ *Ibid*, pp. 45–46.

ment of global soft-law standards for AI governance. In 2019, as one of the first organizations, the OECD published the *OECD Principles on Artificial Intelligence* that are based on values of responsible stewardship of trustworthy AI, which are relevant to all stakeholders. The mentioned document is structured around five value-based principles and five policy recommendations. The value-based principles include: (i) inclusive growth, sustainable development and well-being; (ii) human-centred values and fairness; (iii) transparency and explainability; (iv) robustness, security and safety; and (v) accountability. These principles are complemented by policy recommendations addressed to governments, focusing on areas such as investment in AI research and development, fostering a digital ecosystem for AI, shaping an enabling policy environment, building human capacity, and promoting international cooperation.²²⁵

These principles were revised in 2024, due to rapid technological advancements. The updated version, however, does not represent a departure from the original normative framework adopted in 2019, but rather its contextual refinement. While the 2019 Principles laid the foundation for a value-oriented approach to AI, centred on human-centric values,²²⁶ transparency, accountability, and robustness, the revised 2024 version reflects a noticeable shift towards the operationalisation of these principles in response to the rapid proliferation of advanced AI systems, in particular general-purpose and generative AI.

Importantly, the structural integrity of the original framework has been preserved. The core architecture, consisting of value-based principles and policy recommendations, remains intact, thereby ensuring continuity and normative stability. However, the updated version introduces several substantive clarifications and extensions that respond to challenges which were either marginal or entirely absent at the time of the original adoption.

While the wording of the principle remains largely unchanged, the 2024 update introduces a subtle yet meaningful shift by explicitly incorporating “environmental sustainability” alongside sustainable development and well-

²²⁵ OECD: *Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 22 May 2019, C/MIN(2019)3/FINAL. Online: [https://one.oecd.org/document/C/MIN\(2019\)3/FINAL/en/pdf](https://one.oecd.org/document/C/MIN(2019)3/FINAL/en/pdf).

²²⁶ KLUČKA, J.: *General Overview of the Artificial Intelligence and International Law*. In: KLUČKA, J., BAKOŠOVÁ, L. SISÁK, L. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Nakladatelství Leges, 2021, p. 17.

being.²²⁷ This addition signals a growing recognition of the environmental dimension of AI governance as an autonomous concern, rather than merely an implicit component of broader sustainability objectives. This development is particularly relevant in light of the increasing resource intensity of advanced AI systems, which are associated with significant demands on energy, water, and critical mineral resources, while frequently relying on complex and insufficiently transparent supply chains, the environmental and social impacts of which remain only partially reflected in existing regulatory and reporting frameworks.²²⁸ In this context, such impacts may engage the right to a clean, healthy, and sustainable environment as recognised within the international human rights framework.²²⁹

The 2024 reformulation of the principle reflects a notable shift from an ethics-oriented framing towards a more explicitly legal and governance-based approach. While the 2019 version emphasised “human-centred values,” the updated formulation anchors the principle more firmly in the rule of law, human rights, and democratic values. The OECD framework explicitly anchors AI governance in a wide range of internationally recognised human rights, including non-discrimination and equality, freedom, human dignity, autonomy, and privacy, while also extending its scope to encompass broader normative considerations such as diversity, fairness, social justice, and labour rights. This formulation reflects the inherently multidimensional nature of AI-related risks, which cut across both civil and political rights and socio-economic entitlements. This shift is accompanied by a coarser articulation of risks, including the explicit recognition of misinformation and disinformation amplified by AI, as well as risks arising from unintended or malicious uses of AI systems. Furthermore, the notion of human control is expanded into a broader concept of human agency and oversight,²³⁰ thereby

²²⁷ OECD: *Revised Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 3 May 2024, C/MIN(2024)16/FINAL, para. 1.1. Online: [https://one.oecd.org/document/C/MIN\(2024\)16/FINAL/en/pdf](https://one.oecd.org/document/C/MIN(2024)16/FINAL/en/pdf).

²²⁸ STEIN, A.: *Generative AI and sustainability*. In P. Hacker, P. (ed.): *The Oxford handbook of the foundations and regulation of generative AI*. Oxford: Oxford University Press, 2025. Online: <https://doi.org/10.1093/oxfordhb/9780198940272.001.0001>; HENDERSON, P., HU, J., ROMOFF, J., BRUNSKILL, E., JURAFSKY, D., & PINEAU, J.: *Measuring and mitigating the environmental impact of machine learning*. In: *Journal of Machine Learning Research*, Vol. 21, 2020, pp. 1–44. Online: <https://arxiv.org/abs/2002.05651>.

²²⁹ UN General Assembly: *Right to clean, healthy and sustainable environment*, A/RES/76/300, 28 July 2022, para. 1.

²³⁰ OECD: *Revised Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 3 May 2024, C/MIN(2024)16/FINAL, para. 1.2.

aligning the OECD framework more closely with contemporary regulatory approaches to AI governance.²³¹

While the overall structure of the transparency and explainability principle remains largely unchanged, the 2024 update introduces a more nuanced and operationally grounded formulation. It expands the scope of information to be disclosed by explicitly including data sources, processes, and system limitations, thereby reflecting the increasing complexity of contemporary AI systems, especially generative models. At the same time, the introduction of the qualifier “where feasible and useful” signals a shift towards a more pragmatic approach to explainability, acknowledging the technical constraints associated with certain AI systems. Furthermore, the updated formulation distinguishes more clearly between enabling understanding of AI outputs and ensuring the possibility to challenge them, thus enhancing the functional clarity of the principle.²³²

The 2024 revision of the robustness, security and safety principle reflects a shift from a predominantly preventive and process-oriented approach towards a more operational and intervention-based framework. While the 2019 formulation emphasised traceability and systematic risk management across the AI lifecycle, the updated version introduces specific mechanisms aimed at ensuring control over AI systems in situations of failure or misuse. In particular, the explicit reference to the ability to override, repair, or decommission AI systems signals a move towards a more engineering-oriented conception of AI governance. Furthermore, the updated principle expands its scope by incorporating considerations of information integrity, while simultaneously emphasising the need to safeguard freedom of expression, thereby reflecting the growing societal impact of advanced AI systems.²³³

The accountability principle represents a significant shift from a largely declaratory provision towards a more structured and operational framework. While the 2019 version articulated accountability in general terms, the updated formulation introduces concrete mechanisms through which it is to be realised, including traceability and systematic risk management across the AI system lifecycle. In this respect, accountability is no longer conceived merely as an ex post attribution of responsibility, but rather as a continuous

²³¹ See Chapter 6.3. *The Regional Non-binding Approaches to Artificial Intelligence* and Chapter 7. *Regional Binding Approaches to Artificial Intelligence*.

²³² OECD: *Revised Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 3 May 2024, C/MIN(2024)16/FINAL, para. 1.3.

²³³ *Ibid.*, para. 1.4.

and proactive process embedded within the design, development, deployment, and use of AI systems. Notably, the explicit incorporation of responsible business conduct situates AI governance within a broader framework of corporate responsibility, thereby extending the scope of accountability beyond purely technical considerations. At the same time, the revised principle reflects a broader and more differentiated understanding of risks, explicitly encompassing not only issues related to bias and privacy, but also human rights more generally, as well as labour and intellectual property rights. This expansion signals a move towards a more comprehensive risk taxonomy, aligned with the increasingly complex societal impacts of advanced AI systems. Furthermore, the emphasis on cooperation among multiple actors reflects a shift towards a more distributed model of accountability, acknowledging the interconnected and multi-layered nature of contemporary AI ecosystems. Responsibility is thus no longer confined to a single actor, but is instead shared across a network of stakeholders, including developers, deployers, users, and other entities involved in the AI value chain.²³⁴

In this respect, the updated formulation resonates with the framework established by the *OECD Guidelines for Multinational Enterprises on Responsible Business Conduct*, particularly through its emphasis on risk-based due diligence and the responsibility of business actors to prevent and mitigate adverse impacts arising from their activities. The Guidelines adopt a comprehensive approach to responsible business conduct, extending this responsibility beyond a company's own operations to encompass impacts directly linked to its products, services, or business relationships across the value chain. This is particularly relevant in the context of AI systems, where risks frequently arise within complex and interconnected ecosystems involving multiple actors. Furthermore, the Guidelines cover a broad range of areas, including human rights, labour standards, environmental protection, business integrity, consumer interests, and technological governance, thereby reflecting the inherently cross-cutting nature of risks associated with AI.²³⁵ Regarding the human rights protection, the Guidelines affirm that business enterprises may impact the full spectrum of internationally recognised human rights and are therefore expected to conduct ongoing human rights due diligence. This includes identifying, preventing, mitigating, and accounting for adverse impacts, as well as addressing impacts they cause, contribute to,

²³⁴ *Ibid*, para. 1.5.

²³⁵ OECD: *OECD Guidelines for Multinational Enterprises on Responsible Business Conduct*, OECD Publishing, Paris, 2023. Online: <https://doi.org/10.1787/81f92357-en>.

or are directly linked to through business relationships. The Guidelines emphasise the importance of a risk-based approach, meaningful stakeholder engagement, attention to vulnerable groups, and the establishment of effective remediation mechanisms.²³⁶

Although formally non-binding, their practical significance is reinforced through institutional mechanisms, such as National Contact Points for Responsible Business Conduct, as well as their increasing integration into regulatory and reporting frameworks. As such, they contribute to the consolidation of a governance model in which accountability is not limited to formal legal obligations, but is operationalised through continuous risk management, due diligence processes, and cooperation²³⁷ across actors within the AI value chain.

Building upon this convergence, this evolution is further reinforced by the adoption of the *OECD Due Diligence Guidance for Responsible AI*, which translates the OECD AI Principles into a practical framework for implementation. The Guidance provides structured recommendations for enterprises on how to identify, prevent, mitigate, and account for adverse impacts associated with AI systems, thereby operationalising the concept of responsible business conduct in the context of AI governance. In particular, it establishes due diligence as a central organising principle extending across the entire AI value chain, requiring actors not only to commit to responsible AI, but also to demonstrate how risks—including those related to human rights, labour, and the environment—are effectively managed in practice.²³⁸ These developments illustrate a broader transformation of accountability from a general normative expectation into a functional governance mechanism. The OECD framework thus reflects a gradual shift towards a model of AI governance that combines value-based principles with procedural obligations, aligning soft law instruments more closely with emerging regulatory approaches and reinforcing their role as operational tools within the evolving architecture of global AI governance.

In contrast to the substantive evolution observed in the value-based principles of the OECD Principles on AI, the policy recommendations addressed to governments remain largely unchanged in the 2024 revision. The core areas—investment in AI research and development, fostering an inclusive AI

²³⁶ *Ibid*, pp. 25–27.

²³⁷ *Ibid*, pp. 58–61.

²³⁸ OECD: *OECD Due Diligence Guidance for Responsible AI*, OECD Publishing, Paris, 2026. Online: <https://doi.org/10.1787/41671712-en>.

ecosystem, shaping an enabling and interoperable governance environment, building human capacity, and promoting international co-operation—retain their original structure and content.²³⁹ This continuity suggests that the OECD does not seek to redefine the policy instruments available to governments, but rather to recalibrate their normative orientation through the updated principles. In other words, while the means of AI governance remain stable, the underlying values and priorities guiding their implementation have evolved. This distinction highlights the dual function of the OECD framework: as a relatively stable policy scaffold complemented by a dynamic normative layer capable of adapting to technological and societal developments.

From a broader perspective, the evolution of the OECD AI framework illustrates the increasing sophistication of soft law instruments in the field of AI governance. Rather than merely articulating general principles, the OECD has progressively contributed to their operationalisation, bridging the gap between normative commitments and practical implementation. In doing so, it reinforces the role of soft law not only as a source of guidance, but as a dynamic and adaptive governance mechanism capable of responding to the rapid pace of technological change. The OECD framework thus occupies a pivotal position within the emerging global architecture of AI regulation, functioning as a normative reference point that both reflects and shapes the convergence of international, regional, and domestic approaches to the governance of AI.

6.3. The Regional Non-binding Approaches to Artificial Intelligence

In addition to global soft law initiatives, a growing number of regional organisations have begun to develop non-binding approaches to the governance of AI. These initiatives reflect diverse regional priorities, levels of technological development, and normative traditions, resulting in a plurality of approaches to the regulation of AI and its implications for human rights. While the selected regional frameworks are generally less developed and less formalised than their European counterparts, they play an increasingly important role in shaping regional discourse, setting policy directions, and fos-

²³⁹ OECD: *Revised Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 3 May 2024, C/MIN(2024)16/FINAL, paras. 2.1.–2.5.

tering cooperation among states. At the same time, they contribute to the broader landscape of AI governance by highlighting both converging principles and context-specific regulatory solutions.

This section therefore examines selected regional non-binding approaches to AI, with particular attention to initiatives developed within the AU and the OAS, and their relevance for the protection of human rights and further progressive development of IHRL.

6.3.1. The African Perspective on Artificial Intelligence Regulation

Within the African regional context, the AU has been increasingly involved in shaping the governance of AI. A crucial initial step was taken in 2019, when the Third Ordinary Session of the Specialized Technical Committee on Communication and Specialized Technical Committee on Communication and Information and Communication Technology (STC-CICT), held in Sharm El Sheikh, called for the establishment of a dedicated working group on AI.²⁴⁰ This group was tasked with exploring the development of a common African position on AI, strengthening continental capacity-building efforts, and considering the creation of an AI-focused think tank aligned with the objectives of Agenda 2063²⁴¹ and the Sustainable Development Goals^{242, 243}.

This process was further advanced in 2021, when the ACHPR adopted *Resolution 473 on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa*. The ACHPR frames the governance of AI and related technologies firmly within the existing human rights framework, calling on States to ensure that their development and use remain fully compatible with the *African Charter on Human and Peoples' Rights* and other relevant international

²⁴⁰ AU: *Ministerial Report 3rd Ordinary Session of the African Union Specialized Technical Committee on Communication and Information Technologies (CCICT)*, Sharm El Sheikh, 25–26 October 2019, para. 9. Online: <https://au.int/en/documents/20191025/ministerial-report-3rd-ordinary-session-african-union-specialized-technical>.

²⁴¹ AU: *Agenda 2063: The Africa We Want - Framework Document*, September 2015. Online: https://au.int/sites/default/files/documents/36204-doc-agenda2063_popular_version_en.pdf.

²⁴² UN General Assembly: *Transforming our world: the 2030 Agenda for Sustainable Development*, A/RES/70/1, 21 October 2015.

²⁴³ AU – STC-CICT: *2019 Sharm El Sheikh Declaration*, AU/STC-CICT-3/MIN//Decl., Sharm El Sheikh, 22–26 October 2019, para. 15. Online: https://au.int/sites/default/files/decisions/37590-2019_sharm_el_sheikh_declaration_-_stc-cict-3_oct_2019_ver2410-10pm-1rev-2.pdf.

instruments. In this context, it emphasises a set of core principles, including human dignity, privacy, equality, non-discrimination, inclusion, transparency, accountability and fairness, which should guide all AI-related activities.²⁴⁴ At the same time, the Commission highlights the need to adapt AI systems to African contexts, stressing the importance of incorporating local values, norms and societal needs in order to address broader concerns of epistemic injustice and external technological dependence. It further underlines the centrality of transparency and explainability, requiring that AI-driven decisions be understandable to affected individuals.²⁴⁵ From a regulatory perspective, the Commission calls for the development of comprehensive legal and ethical governance frameworks at both national and regional levels, including stronger engagement by the AU and regional bodies.²⁴⁶ A particularly significant element is the insistence on maintaining meaningful human control over AI systems, which is framed as an emerging human rights principle aimed at preventing adverse impacts on fundamental rights.²⁴⁷ Finally, the ACHPR signals the need for continued normative development in this field, committing to further research and the elaboration of guidelines in co-operation with expert bodies, thereby contributing to the progressive development of AI-related human rights standards in the African context.²⁴⁸

The development of the AU's *Continental Artificial Intelligence Strategy* was preceded by the adoption of a *Conceptual Framework on AI*, approved in 2023 by the STC-CICT and subsequently endorsed by the AU Executive Council. On this basis, the AU Commission was mandated to develop a comprehensive and forward-looking strategy aimed at harnessing the transformative potential of AI in line with *Agenda 2063*,²⁴⁹ while addressing its

²⁴⁴ ACHPR: *Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa*, ACHPR/Res. 473 (EXT.OS/ XXXI) 2021, 25 February 2021, para. 1. Online: <https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-human-and-peoples-rights-and-art>.

²⁴⁵ *Ibid*, para. 2–3.

²⁴⁶ *Ibid*, para. 4–5.

²⁴⁷ *Ibid*, para. 6; CAVALCANTE SIEBERT, L., LUPETTI, M. L., AIZENBERG, E. et al.: *Meaningful human control: actionable properties for AI system development*. In: *AI Ethics*, Vol. 3, 2023, pp. 241–255.

²⁴⁸ ACHPR: *Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa*, ACHPR/Res. 473 (EXT.OS/ XXXI) 2021, 25 February 2021, para. 7.

²⁴⁹ AU: *Agenda 2063: The Africa We Want – Framework Document*, September 2015. Online: https://au.int/sites/default/files/documents/33126-doc-framework_document_book.pdf.

associated risks. The strategy was further shaped by multi-stakeholder consultations conducted in 2024, focusing on key areas such as maximising the benefits of AI, mitigating risks, strengthening capabilities and fostering co-operation. Its formulation also drew on existing AU policy frameworks and was supported by a broad network of institutional actors, including AU bodies, regional organisations and international partners such as UNESCO and United Nations Economic Commission for Africa, thereby ensuring that it reflects a wide range of perspectives and continental priorities.²⁵⁰

Against this background, the *Continental Artificial Intelligence Strategy* articulates a set of guiding principles intended to shape the development and governance of AI across the continent in a manner responsive to Africa's specific social, economic, and developmental realities. Central among these is the “local-first” principle, which prioritises the development and deployment of AI solutions tailored to African challenges, while promoting local talent, innovation ecosystems, and the preservation of cultural values. This is closely linked to a people-centred approach, under which AI is expected to contribute to inclusive growth, sustainable development, and the well-being of communities, particularly in underserved and rural areas.

Central to the Strategy is the explicit commitment to human rights and human dignity, requiring that all AI systems be developed and used in accordance with the African *Charter on Human and Peoples' Rights* and broader IHRL. These principles are complemented by a strong emphasis on inclusion and diversity, ensuring that AI systems are non-discriminatory and responsive to the continent's social, cultural and linguistic plurality, while also promoting gender equality and the empowerment of vulnerable groups. The Strategy further incorporates principles of ethics and transparency, aimed at preventing bias, inequality and social exclusion, as well as safeguarding cultural identity and mitigating technological divides. In parallel, it highlights the importance of peace and prosperity, linking AI governance to broader objectives of societal stability, security and environmental protection. Finally, the principles of cooperation, integration and skills development underscore the need for regional collaboration, data sovereignty and capacity-building, including through education and public awareness, as essential components of a sustainable and inclusive AI ecosystem in Africa.²⁵¹

²⁵⁰ AU: *Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity*, 2024, p. 16. Online: https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf.

²⁵¹ *Ibid.*, p. 28.

The Strategy is operationalised through five interconnected focus areas and fifteen policy actions aimed at linking AI governance with broader socio-economic and developmental objectives under Agenda 2063. Particular emphasis is placed on the responsible use of AI in sectors such as healthcare, education, agriculture, public administration, and climate action, while simultaneously addressing risks relating to discrimination, information integrity, technological inequality, and social exclusion.

From a human rights perspective, the Strategy promotes the adoption of governance and regulatory frameworks at both national and regional levels, alongside ethical principles and technical standards intended to ensure that AI systems remain safe, transparent, inclusive, and consistent with regional and international human rights standards. At the same time, it seeks to strengthen institutional and technological capacities across the continent through investment in digital infrastructure, research, innovation, AI-related skills, and multi-stakeholder cooperation involving governments, academia, civil society, and the private sector.²⁵²

Subsequently, during the Global AI Summit on Africa in 2025, States adopted the *Africa Declaration on Artificial Intelligence* in order to articulate a shared continental vision for AI governance grounded in African priorities and values. From a human rights perspective, the Declaration emphasises the need to implement safeguards preventing harm and protecting privacy, ethics, transparency, and explainability, while prioritising human dignity, fundamental rights, freedoms, and environmental sustainability in the development and deployment of AI systems.²⁵³

These developments build upon earlier regional instruments, particularly the *AU Data Policy Framework*²⁵⁴ and the *Malabo Convention on Cyber Security and Personal Data Protection*,²⁵⁵ which together provide important normative foundations for data governance, cybersecurity, digital rights, and cross-border digital cooperation within the African context. However, they

²⁵² *Ibid*, pp. 4–5.

²⁵³ *Africa Declaration on Artificial Intelligence*, Global AI Summit on Artificial Intelligence, 4 April 2025, para. 2.1.2. Online: <https://c4ir.rw/docs/Africa-Declaration-on-Artificial-Intelligence.pdf>.

²⁵⁴ AU: *AU Data Policy Framework*, 28 July 2022. Online: <https://au.int/sites/default/files/documents/42078-doc-DATA-POLICY-FRAMEWORKS-2024-ENG-V2.pdf>.

²⁵⁵ *African Union Convention on Cyber Security and Personal Data Protection*, Malabo, 27 June 2014.

gesture towards innovation and development, but largely sidestep the deeper structural risks AI poses to dignity, equality, and accountability.²⁵⁶

Collectively, these developments demonstrate that the emerging African approach does not merely replicate global AI governance models, but increasingly seeks to adapt them to the continent's specific human rights, developmental and socio-economic context. Its particular contribution lies in combining universal human rights standards with region-specific concerns relating to technological dependence, cultural and linguistic diversity, digital inequality and capacity-building.²⁵⁷ At the same time, the predominantly strategic and non-binding character of these instruments means that their effectiveness will depend largely on their translation into national and regional regulatory frameworks and institutional practice.

6.3.2. The Organization of American States

Within the Inter-American system, the regulation of AI has developed primarily through soft law initiatives, policy cooperation, and thematic discussions concerning the human rights implications of emerging digital technologies. Although the OAS has not adopted a comprehensive binding framework specifically dedicated to AI, its institutions have increasingly engaged with issues relating to digital governance, algorithmic decision-making, data protection, and the societal impacts of AI systems.

Relevant in this regard is the Inter-American Framework for Data and AI Governance, developed through a regional multi-stakeholder process facilitated by the OAS Department for Effective Public Management. The framework was prepared by an ad hoc working group on data governance and AI composed of government representatives participating in the Inter-American Network on Digital Government. It recognises that effective data governance constitutes a necessary precondition for trustworthy and effective AI governance and emphasises the importance of developing coherent regulatory and institutional frameworks capable of supporting the responsible deployment of AI technologies. In this respect, the framework seeks to promote the use of AI to improve public service delivery, reduce inequalities,

²⁵⁶ FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: *African Journal of Legal Research*, Vol. 4, No. 2, 2026, p. 55.

²⁵⁷ See NCUBE, C., ORIAKHOGBA, D., RUTENBERG, I., SCHONWETTER, T. (eds.): *Artificial Intelligence and the Law in Africa*, LexisNexis, 2023, p. 71. Online: <https://cipit.strathmore.edu/artificial-intelligence-and-the-law-in-africa/>.

create economic opportunities, and support inclusive and sustainable development across the region.²⁵⁸

The framework is structured around three principal components. First, it contains the *Inter-American guidelines for data and artificial intelligence governance*, which was formulated through a collaborative regional process conducted in 2023. During this process, participating States identified twelve priority thematic areas, which included, *inter alia*, institutional capacity-building, alignment with international initiatives, ethical and human rights safeguards, data governance structures, digital infrastructure, cybersecurity, open data, research and innovation, and multi-stakeholder cooperation. Particular emphasis is placed on transparency, human rights protection, data governance culture, and the development of inclusive ecosystems capable of maximising the societal and economic benefits of AI and data-driven technologies. These thematic areas intended to provide a flexible roadmap for the development of coherent and rights-oriented approaches to data and AI governance within the region.

From the human rights perspective, it encourages alignment with UNESCO Recommendations and relevant UN initiatives,²⁵⁹ while promoting safeguards relating to privacy, personal data protection, transparency, explainability, and open government principles.²⁶⁰ Particular attention is devoted to risk-based governance approaches, including impact assessments, audits, and mitigation measures aimed at preventing adverse effects on individual rights.²⁶¹ The framework further advocates inclusive and multi-stakeholder oversight mechanisms involving civil society, academia, and the private sector, alongside effective grievance and remedy procedures for rights violations arising from the use of AI and data-driven technologies. It also promotes culturally sensitive and inclusive AI governance models attentive to the needs of vulnerable groups and regional diversity.²⁶²

The Framework is also accompanied by a regional study assessing the current state of data and AI governance across the Americas and a model policy instrument designed to assist Member States in developing or adapting

²⁵⁸ OAS: *Inter-American Guidelines on Data Governance and AI*, 2024, p. 4. Online: <https://www.oas.org/ext/DesktopModules/MVC/OASDnnModules/Views/Item/Download.aspx?type=1&id=1080&lang=1>.

²⁵⁹ UN General Assembly: *Seizing the Opportunities of Safe, Secure, and Trustworthy Artificial Intelligence Systems for Sustainable Development*, A/RES/78/265, 21 March 2024.

²⁶⁰ OAS: *Inter-American Guidelines on Data Governance and AI*, 2024, para. 5.1.–5.5.

²⁶¹ *Ibid*, para. 5.6.

²⁶² *Ibid*, para. 5.7.–5.10.

domestic governance frameworks relating to data governance and AI regulation. However, while these components are referenced within the framework, some accompanying materials do not currently appear to be publicly accessible in full detail.²⁶³ Collectively, these tools are intended to support OAS Member States and local authorities in strengthening institutional capacities, promoting regulatory harmonisation, and fostering the development of trustworthy, transparent, and human rights-oriented governance approaches to emerging digital technologies. The Inter-American approach thus reflects a gradual movement from general digital governance cooperation towards a more structured rights-oriented framework for AI. Its reliance on existing human rights standards, risk-based governance and institutional capacity-building demonstrates the adaptability of the regional framework to technological change. Nevertheless, the absence of binding AI-specific obligations continues to limit the degree of regulatory convergence and enforceability across the region.

²⁶³ OAS: Inter-American Framework on Data Governance and Artificial Intelligence (MIG-DIA). Online: <https://www.oas.org/ext/en/democracy/inter-american-framework-on-data-governance-and-artificial-intelligence-migdia>.

7. REGIONAL BINDING APPROACHES TO ARTIFICIAL INTELLIGENCE

In contrast to the predominantly non-binding approaches at the global level, certain regional actors have begun to develop binding regulatory frameworks for AI. Most notably, the EU and the COE have adopted legal instruments that establish enforceable standards within their respective jurisdictions, thereby marking a significant shift from principle-based governance towards more structured legal regulation. Against this background, this chapter examines the emerging regional binding approaches to AI, with a particular focus on the European legal context. It analyses the regulatory framework developed by the EU, including the *Artificial Intelligence Act*, as well as the COE's *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*. Attention is paid to the extent to which these instruments integrate human rights considerations into the governance of AI. At the same time, the analysis also considers earlier normative developments at the regional level, including a range of non-binding instruments adopted prior to the AI Act and the Framework Convention. Although lacking formal legal force, these instruments have played an important role in shaping the conceptual foundations, guiding principles and regulatory architecture of subsequent binding frameworks. Their examination therefore allows for a more comprehensive understanding of the normative evolution of AI governance in Europe and the interplay between soft law and binding regulation. Moreover, both instruments exert a broader normative influence beyond the European region, shaping regulatory approaches, policy discussions, and standard-setting efforts in other parts of the world. The central question is therefore whether the transition from non-binding principles to legally enforceable obligations is capable of addressing the implementation and accountability gaps identified in the preceding Chapter, while preserving sufficient flexibility to respond to rapidly evolving AI technologies.

7.1. The European Union's Regulatory Framework for Artificial Intelligence

The development of the EU framework on AI illustrates the gradual shift from ethical²⁶⁴ and policy-oriented discussions²⁶⁵ towards increasingly detailed legal regulation of AI systems. In this process, the protection of fundamental rights emerged as one of the central organising principles of the evolving regulatory architecture, influencing both the substantive obligations imposed on AI providers and the institutional mechanisms responsible for supervision and enforcement, while simultaneously seeking to support technological innovation, economic competitiveness, and the development of a trustworthy European AI ecosystem.²⁶⁶

Before examining the *Artificial Intelligence Act* (hereinafter “AI Act”) in greater detail, it should be noted that the EU legal framework already includes several binding instruments relevant to the protection of fundamental rights in the context of AI systems. The normative foundations of the EU approach to AI regulation are rooted in the values enshrined in Art. 2 of the *Treaty on European Union*, according to which the Union is founded on respect for human dignity, freedom, democracy, equality, the rule of law, and respect for human rights. These values constitute the broader constitutional framework within which the EU's regulatory response to AI has gradually evolved. At its core lies the *Charter of Fundamental Rights of the European*

²⁶⁴ High-Level Expert Group on Artificial Intelligence: *Ethics Guidelines for Trustworthy AI*, Brussels, 2019. Online: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

²⁶⁵ See e.g. *Declaration on Artificial Intelligence Cooperation*, 10 April 2018, pp. 3–5. Online: <https://ec.europa.eu/jrc/communities/en/community/digitranscope/document/eu-declaration-cooperation-artificial-intelligence>; European Commission: *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe*, 25 April 2018, COM(2018)/237; European Commission: *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence*, 7 December 2018, COM(2018) 795; European Commission: *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Fostering a European Approach to Artificial Intelligence*, Brussels, 21 April 2021, COM(2021) 205 final.

²⁶⁶ European Commission: *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Fostering a European Approach to Artificial Intelligence*, Brussels, 21 April 2021, COM(2021) 205 final, p. 1 and 4.

Union, which provides a comprehensive catalogue of human rights applicable within the scope of EU law, including those particularly relevant to the deployment of AI, such as the rights to private and family life, protection of personal data, non-discrimination, and effective remedy and to a fair trial.²⁶⁷

Well before the adoption of dedicated AI legislation, the EU had already developed an extensive regulatory landscape capable of addressing many of the risks associated with digital technologies. This framework is anchored in cornerstone instruments such as the *General Data Protection Regulation*²⁶⁸ in the field of privacy and data protection, alongside a broader set of digital governance measures, including the *Data Governance Act*,²⁶⁹ the *Digital Services Act*,²⁷⁰ and the *Digital Markets Act*.²⁷¹ These instruments collectively seek to ensure transparency, accountability, fair competition, and the responsible use of data-driven technologies, while balancing innovation with the protection of fundamental rights and societal interests.²⁷² The growing importance of these instruments is closely linked to the data-intensive nature of contemporary AI systems, particularly those based on machine learning and big data analytics. In this respect, the GDPR has become particularly relevant, as many AI systems rely on personal data both for training algorithmic models and for automated decision-making affecting individuals. Although the GDPR does not explicitly regulate AI, its principles—including purpose limitation, data minimisation, transparency, accountability, and restrictions on automated decision-making—have become central to the gov-

²⁶⁷ *Charter of Fundamental Rights of the European Union*, 26.10.2012, OJ C 326, pp. 391–407, Art. 6, 8, 21, 47.

²⁶⁸ *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC*, OJ L 119/1.

²⁶⁹ *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)*, OJ L 152.

²⁷⁰ *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)*, OJ L 277.

²⁷¹ *Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)*, OJ L 265.

²⁷² See e.g. *Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act)*, OJ L 2023/2854, Arts. 1, 4 and 6, Chapter III.

ernance of AI systems within the EU legal order.²⁷³ At the same time, the application of these principles to AI technologies has generated important interpretative and practical challenges, particularly regarding profiling, automated decision-making, explainability, and the compatibility of large-scale data processing with data protection requirements.

From the perspective of AI governance, the GDPR grants data subjects several important rights aimed at strengthening transparency, accountability, and individual control over personal data processing. In particular, Art. 15 GDPR requires controllers to provide information about automated decision-making, including meaningful information about the logic involved and the potential consequences of such processing.²⁷⁴ Although the precise scope of this obligation remains debated, especially regarding a possible “right to explanation”, it nevertheless constitutes an important transparency safeguard applicable to AI systems. Additional protections are provided through the rights to erasure, portability, and objection, particularly in relation to profiling and automated processing. Collectively, these rights reflect the effort to reconcile AI-driven data processing with the protection of privacy, autonomy, and informational self-determination.²⁷⁵ At the same time, the limitations of existing data protection and digital governance rules gradually contributed to the emergence of AI-specific regulation within the EU legal order.

In parallel, EU product safety legislation has also progressively adapted to the risks associated with digital technologies and AI-integrated products. This development is reflected in instruments such as the *Machinery Directive*, now being replaced by the *Machinery Regulation*,²⁷⁶ and the *General Product*

²⁷³ SARTOR, G., LAGIOIA, F.: *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, STOA | Panel for the Future of Science and Technology, 2020, pp. 44–48. Online: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU\(2020\)641530_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf).

²⁷⁴ *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC*, OJ L 119/1, Art. 15.

²⁷⁵ SARTOR, G., LAGIOIA, F.: *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, STOA | Panel for the Future of Science and Technology, 2020, pp. 56–58.

²⁷⁶ *Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on machinery and repealing Directive 2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC*, OJ L 165. The Directive 2006/42/EC will be replaced from 20 January 2027.

Safety Directive, superseded by the *General Product Safety Regulation*.²⁷⁷ The Machinery Regulation in particular explicitly addresses safety risks arising from AI-enabled products by requiring appropriate risk assessments, conformity assessments, and safety requirements throughout the product life-cycle.²⁷⁸

Additional layers of protection are provided by instruments such as the *Cybersecurity Act*,²⁷⁹ the *Consumer Rights Directive*,²⁸⁰ and a broader set of regulatory measures, including the *Product Liability Directive*. For instance, the revised *Product Liability Directive* further adapts the EU liability framework to the realities of digital technologies by expressly extending the notion of a “product” to software, including software incorporating AI systems. At the same time, the Directive reflects that the assessment of defectiveness must take into account factors such as the ability of products to continue learning or acquiring new features after being placed on the market, the impact of software updates and upgrades, and compliance with safety-relevant cybersecurity requirements.²⁸¹ The Directive further acknowledges that economic operators²⁸² may retain control over digital products even after their release on the market, particularly through software updates, related services, or substantial modifications. Consequently, liability may continue to

²⁷⁷ Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC, OJ L 135.

²⁷⁸ Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on machinery and repealing Directive 2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC, OJ L 165, recital 32, 55, Art. 10.

²⁷⁹ Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013, OJ L 151.

²⁸⁰ Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council Text with EEA relevance, OJ L 304.

²⁸¹ Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC, OJ L, 2024/2853, 18.11.2024, Art. 7.

²⁸² Economic operator means a manufacturer of a product or component, a provider of a related service, an authorised representative, an importer, a fulfilment service provider or a distributor. *Ibid*, Art. 4 (15).

arise where economic operators fail to implement necessary updates or security measures required to address evolving technological and cybersecurity risks.²⁸³ In this way, the Directive strengthens accountability mechanisms for harm caused by AI-enabled products and better aligns the liability framework with the continuous, adaptive, and data-driven character of contemporary digital technologies.

Taken together, these instruments form a dense and interconnected regulatory ecosystem that provides a baseline of protection and within which AI systems are developed and deployed. However, the regulatory challenges associated with the development and use of AI prompted the EU to develop a more comprehensive regulatory approach, combining the adaptation of existing legal instruments with the introduction of new legislative measures specifically addressing AI. Recent developments in areas such as product liability, cybersecurity, and data governance illustrate an effort to recalibrate the legal framework considering technological change, strengthening accountability and risk management mechanisms.

These developments culminated in the adoption of the *Artificial Intelligence Act*, which represents the first comprehensive attempt to establish a harmonised and binding regulatory framework for AI within the EU. The following section therefore examines the evolution of the EU's approach to AI governance and analyses the key features of the AI Act, with particular attention to its implications for the protection of human rights.

7.1.1. The Artificial Intelligence Act and the Protection of Human Rights

The AI Act was formally approved by the Council in May 2024 and represents the first comprehensive binding regulatory framework specifically dedicated to AI at the supranational level. Its adoption reflects several years of EU efforts aimed at developing a coherent approach to AI governance, gradually evolving from policy strategies and coordinated action plans into a legally binding regulatory instrument.²⁸⁴

²⁸³ *Ibid*, Art. 11.

²⁸⁴ For a more detailed analysis of the preparatory work relating to the AI Act, see BAKOŠOVÁ, L.: *Green Ambitions in the Era of Artificial Intelligence*. Praha: Nakladatelství Leges, 2024, pp. 139–141; STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, pp. 46–50.

The AI Act seeks to improve the functioning of the internal market while promoting the development and uptake of human-centric and trustworthy AI. At the same time, it aims to ensure a high level of protection of health, safety, fundamental rights, democracy, the rule of law, and environmental protection against the potentially harmful effects of AI systems, while simultaneously supporting technological innovation. To achieve these objectives, the Regulation establishes a harmonised legal framework governing the placing on the market, deployment, and use of AI systems within the EU. It introduces prohibitions of certain unacceptable AI practices, sets specific requirements and obligations for high-risk AI systems and their operators, and imposes transparency obligations for selected categories of AI systems. In addition, the AI Act regulates general-purpose AI models, establishes rules relating to governance, market surveillance, and enforcement, and incorporates measures intended to support innovation, particularly regarding small and medium-sized enterprises (hereinafter “SMEs”) and start-ups.²⁸⁵

It applies to a wide range of actors involved in the development, deployment, distribution, and use of AI systems, thereby ensuring a comprehensive governance framework extending both within and beyond the EU. Its scope encompasses providers placing AI systems or general-purpose AI models on the EU market, deployers operating within the Union, importers, distributors, authorised representatives, and manufacturers integrating AI systems into products marketed under their own name or trademark. The Regulation also applies, under certain conditions, to providers and deployers established outside the EU where the outputs of AI systems are used within the Union.²⁸⁶ In this way, the AI Act establishes a broad framework of accountability and regulatory oversight across the AI lifecycle while simultaneously reflecting the EU’s broader ambition to influence the development of global standards for AI governance.²⁸⁷

At the centre of the AI Act lies a risk-based regulatory model under which AI systems are categorised according to the level of risk they pose to fundamental rights, safety, and other protected interests. This approach is mo-

²⁸⁵ AI Act, Art. 1.

²⁸⁶ *Ibid*, Art. 2 (1).

²⁸⁷ CANCELA-OUTEDA, C.: *The EU’s AI act: A framework for collaborative governance*. In: Internet of Things, Vol. 27, 2024, p. 6; MELTZER, J., TIELEMANS, A.: *The European Union AI Act: Next steps and issues for building international cooperation*, Global Economy and Development at Brookings, May 2022, p. 1. Online: <https://www.brookings.edu/research/the-european-union-ai-act-next-steps-and-issues-for-building-international-cooperation-in-ai/>

tivated by the need to find a balance between encouraging innovation and ensuring AI systems function safely, transparently, and in congruence with fundamental rights and European values.²⁸⁸ The Regulation consequently imposes graduated obligations on actors involved in the AI lifecycle, with the strictest requirements applying to high-risk AI systems. These include obligations relating to risk management, data governance, transparency, human oversight, and conformity assessment procedures prior to the placing of such systems on the market.²⁸⁹

The scope of the AI Act is nevertheless subject to several important exclusions. It does not apply to AI systems developed exclusively for scientific research and development unless they are tested in real-world conditions, nor to AI systems used solely for personal and non-professional purposes. Certain open-source AI systems are likewise excluded, except where they qualify as high-risk. In addition, the AI Act does not govern activities relating to military, defence, or national security, and contains limited exemptions for specific activities carried out by public authorities of third countries and international organisations in the context of law enforcement and judicial cooperation.²⁹⁰ These exclusions reflect the EU's attempt to reconcile the objectives of AI regulation with broader security interests, existing legal frameworks, and the need to avoid disproportionate regulatory burdens in low-risk contexts.

The AI Act establishes a comprehensive framework of prohibited and regulated AI practices aimed at safeguarding fundamental rights and freedoms within the EU legal order. In particular, it prohibits AI systems employing manipulative or deceptive techniques capable of distorting individual behaviour, systems exploiting vulnerabilities linked to characteristics such as age or disability, and social scoring practices resulting in unjustified or disproportionate treatment of individuals or groups. The Regulation also restricts certain forms of biometric surveillance and profiling, including AI systems used to predict criminal behaviour based solely on profiling or personality traits, as well as the untargeted scraping of facial images for the creation of facial recognition databases. Particular attention is devoted to the use of remote

²⁸⁸ STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, p. 52.

²⁸⁹ *AI Act*, Art. 9–15.

²⁹⁰ *Ibid.*, Art. 2 (3)–(12).

biometric identification technologies in publicly accessible spaces for law enforcement purposes. Such systems may only be used under narrowly defined circumstances, including the search for missing persons or the prevention of imminent threats to public security, and remain subject to strict safeguards such as prior authorisation, necessity and proportionality requirements, and fundamental rights impact assessments. The Regulation further limits the use of emotion recognition systems in workplaces and educational institutions, except where justified by medical or safety considerations. Through these prohibitions and safeguards, the AI Act seeks to balance technological innovation with the protection of human dignity, privacy, autonomy, and other fundamental rights affected by the deployment of AI systems.²⁹¹

The central focus of the AI Act lies in the regulation of high-risk AI systems, which are subject to the most stringent obligations under the Regulation. These include AI systems used as safety components of products covered by harmonised EU legislation,²⁹² as well as specific use cases listed in Annex III due to their potential impact on fundamental rights and safety in areas such as biometrics, critical infrastructure, education, employment, access to essential public and private services, law enforcement, migration, and the administration of justice. Examples include biometric identification systems, AI systems used in recruitment and employment decisions, creditworthiness assessments, migration and asylum procedures, and systems supporting judicial or law enforcement authorities in decision-making processes. At the same time, systems performing only limited procedural or supportive functions are excluded from the high-risk category.²⁹³ The classification of high-risk systems remains flexible, as the Commission may amend the list of high-risk use cases in response to technological developments and emerging risks. Providers of non-high-risk systems are nevertheless required to document their assessment before placing such systems on the market, while the Commission is expected to issue further guidance clarifying the classification framework.²⁹⁴

²⁹¹ *Ibid*, Art. 5.

²⁹² For instance, *Directive 2006/42/EC on machinery*; *Directive 2009/48/EC on the safety of toys*; *Regulation (EU) 2017/745 on medical devices*; *Regulation (EU) 2016/424 on cableway installations*; *Directive (EU) 2016/797 on the interoperability of the rail system within the European Union*; *Regulation (EU) No 167/2013 on the approval and market surveillance of agricultural and forestry vehicles*.

²⁹³ *AI Act*, Art. 6 (2).

²⁹⁴ *Ibid*, Art. 6–7.

These systems undergo rigorous regulatory scrutiny, necessitating thorough risk assessment, high-quality datasets to reduce bias, and providing transparency norms, user information provisions, robustness checks, accuracy standards, and human monitoring.²⁹⁵ By contrast, AI systems considered to present only limited risks are subject to less stringent obligations, primarily focused on transparency requirements. For example, users interacting with AI systems such as chatbots must be informed that they are communicating with a machine.²⁹⁶

The AI Act establishes a distinct regulatory regime for general-purpose AI (hereinafter „GPAI“) models,²⁹⁷ reflecting their broad applicability and potential systemic impact across multiple downstream sectors and applications. Providers of GPAI models are subject to a range of obligations aimed at ensuring transparency, accountability, and compliance with EU law. These include requirements to prepare and maintain technical documentation, provide relevant information to downstream AI system providers, comply with EU copyright rules, and publish sufficiently detailed summaries concerning the data used for model training. At the same time, providers are expected to cooperate with supervisory authorities and may rely on codes of practice and harmonised standards to demonstrate compliance.²⁹⁸ Certain open-source GPAI models remain exempt from some of these obligations unless they present systemic risks. Additional and significantly stricter obligations apply to GPAI models classified as presenting systemic risks. Providers of such models are required to conduct model evaluations and adversarial testing, identify and mitigate systemic risks at the EU level, report serious incidents to competent authorities, and ensure a high level of cybersecurity protection. The AI Act further envisages the development of EU-level codes of practice coordinated by the AI Office, intended to facilitate consistent implementation of obligations relating to transparency, risk management, and account-

²⁹⁵ *Ibid*, Art. 8–15; STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, p. 53.

²⁹⁶ *Ibid*, pp. 53–54.

²⁹⁷ An AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market. *AI Act*, Art. 3 (63).

²⁹⁸ *Ibid*, Art. 53.

ability. Through this framework, the Regulation seeks to address the unique challenges posed by increasingly powerful and widely deployable AI models whose impacts may extend across sectors, jurisdictions, and fundamental rights interests.²⁹⁹ Collectively, these measures seek to strengthen public trust in AI technologies while ensuring that high-risk AI systems remain consistent with EU standards relating to safety, accountability, transparency, and the protection of fundamental rights.

A particularly important human rights safeguard introduced by the AI Act is the obligation under Art. 27 requiring deployers of certain high-risk AI systems to carry out a fundamental rights impact assessment (hereinafter “FRIA”) prior to deployment. This mechanism reflects the broader rights-based orientation of the Regulation by seeking to ensure that the use of high-risk AI systems is assessed not only from a technical or safety perspective, but also regarding its potential implications for individuals and affected groups. The assessment must examine, *inter alia*, the intended use of the system, the categories of persons likely to be affected, the potential risks to fundamental rights, and the safeguards implemented to mitigate such risks, including measures relating to human oversight, governance, and complaint procedures.³⁰⁰ Once the assessment has been performed, the deployer will generally have to notify the market surveillance authority of its results. In this respect, the FRIA constitutes a preventive accountability mechanism aimed at identifying and addressing risks relating to discrimination, privacy, autonomy, due process, and other protected rights before the AI system is deployed in practice. The mechanism further complements existing obligations under EU data protection law, particularly data protection impact assessments under the GDPR, thereby reinforcing the broader integration of fundamental rights considerations within the EU framework for AI governance. At the same time, doubts have been raised as to whether the AI Act can genuinely be regarded as a “gold standard” in the protection of fundamental rights. Critics have pointed to the limited scope of the FRIA obligation, the absence of a clear duty to prevent unacceptable risks before deployment, and the broad exemptions applicable to law enforcement and migration authorities. These limitations illustrate the broader tensions within the AI Act between the protection of fundamental rights and competing security and

²⁹⁹ *Ibid*, Art. 55–56.

³⁰⁰ *Ibid*, Art. 27.

operational considerations.³⁰¹ At the same time, the Regulation attempts to address these concerns through a broader framework of supervision, enforcement, and accountability mechanisms. To this end, Member States are required to designate competent national authorities responsible for monitoring compliance with the Regulation, including market surveillance activities. The AI Act further provides for administrative sanctions and fines in cases of non-compliance, while leaving the detailed enforcement mechanisms and applicable penalties to national legal systems.³⁰² Through this supervisory framework, the Regulation seeks to strengthen accountability and promote the responsible development and use of AI technologies across the EU.³⁰³ The supervisory mechanism established under the AI Act is examined in greater detail in *Chapter 9.1.1. Governance Architecture under the Artificial Intelligence Act*

The broader EU debate on AI governance has also extended to the field of civil liability and access to remedies for harm caused by AI systems. In this context, the European Commission proposed the *AI Liability Directive*, intended to harmonise certain aspects of national non-contractual liability rules and facilitate compensation claims relating to damage caused by AI systems. The proposal aimed to ensure that individuals harmed by AI systems would benefit from a level of protection comparable to situations in which AI technologies are not involved, thereby strengthening coherence with the broader framework established by the AI Act. In particular, the proposal sought to alleviate evidentiary difficulties typically associated with complex and opaque AI systems, including through mechanisms allowing courts to order the disclosure of relevant evidence concerning high-risk AI systems and through rebuttable presumptions of causality where proving the causal link would otherwise be excessively burdensome for the claimant.³⁰⁴ Although the proposal was subsequently withdrawn, it nevertheless

³⁰¹ PUSZTAHELYI, R.: *Some Reflections on Fundamental Rights Impact Assessment for High-Risk AI Systems (FRIA)*. In: *The Juridical Current Journal*, Vol. 103, No. 4, 2025, pp. 1384–1385.

³⁰² *AI Act*, Art. 99–101.

³⁰³ STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, p. 60.

³⁰⁴ *Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive)*, 28. 9. 2022, COM/2022/496 final, Art. 2–3.

illustrates the broader recognition within the EU legal order that effective AI governance requires not only ex ante regulatory obligations, but also adequate mechanisms of accountability and access to remedies for individuals affected by AI-related harms.

Overall, the AI Act represents a significant shift from principle-based AI governance towards legally enforceable, risk-based regulation. Its principal contribution to fundamental rights protection lies in translating general rights and regulatory principles into preventive obligations concerning risk management, transparency, human oversight and impact assessment. However, the Regulation does not constitute a self-contained human rights framework. Its effectiveness depends on its interaction with the Charter, the GDPR and other existing EU legal instruments, as well as on effective supervision, enforcement and access to remedies. The EU approach therefore illustrates a complementary model in which AI-specific regulation strengthens, rather than replaces, the existing system of fundamental rights protection.

7.2. The Council of Europe's Regulation on Artificial Intelligence

The COE, similarly to the EU, initially focused on the identification of ethical principles in the early stages of AI regulation. These principles were primarily articulated through non-binding instruments, which laid the foundations for the subsequent development of regulatory strategies and represented some of the earliest thematic responses of the COE to technological advancements in specific areas of social life. Such instruments include the *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the impacts of algorithmic systems on human rights, the Commissioner for Human Rights' recommendation "Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights", and Resolution No. 2341 (2020) of the Parliamentary Assembly of the Council of Europe on the need for democratic governance of artificial intelligence. However, the most significant instrument is the *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, which represents a shift from soft law approaches towards a structured and enforceable regulatory framework within the COE system, even though its scope extends beyond the organisation itself, as it has been opened to and joined by States that are not members of the COE.

7.2.1. Normative Foundations of the Framework Convention: The Role of Soft Law

To fully understand the normative foundations of the Framework Convention, it is necessary to examine the preceding soft law instruments developed within the COE. These instruments not only articulated the initial ethical and human rights principles relevant to AI but also contributed to shaping the conceptual and regulatory architecture later reflected in the binding framework. The following section therefore analyses selected pre-Convention instruments, with a view to identifying the key normative trends and continuities that informed the transition from principle-based guidance to legally binding regulation.

One of the earliest and most significant instruments in this field is the *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, adopted by the European Commission for the Efficiency of Justice in 2018.³⁰⁵ The Charter establishes five core ethical principles intended to guide policymakers, legislators and judicial actors in responding to the growing use of AI in judicial contexts. At its core lies the principle of respect for fundamental rights, which requires that the design and deployment of AI systems in judicial settings be fully compliant with the rights guaranteed under the ECHR and the relevant data protection instruments.³⁰⁶ In particular, the use of AI must not undermine the right of access to a court or the right to a fair trial.³⁰⁷ Closely linked to this is the principle of non-discrimination, which emphasises the need to prevent biases and unequal treatment, especially where AI systems rely on sensitive data such as racial or ethnic origin, political opinions, health or other personal characteristics. Where risks of discrimination arise, appropriate corrective measures must be adopted. The Charter further highlights the importance of quality, safety and transparency, requiring the use of reliable data and secure

³⁰⁵ European Commission for the Efficiency of Justice: *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, Council of Europe: Strasbourg, 3–4 December 2018. Online: <https://rm.COE.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

³⁰⁶ *The Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*, ETS No. 108, Strasbourg, 28 January 1981; *Additional protocol to Convention 108 regarding supervisory authorities and transborder data flows*, ETS No. 181, Strasbourg, 8 November 2001.

³⁰⁷ European Commission for the Efficiency of Justice: *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, Council of Europe: Strasbourg, 3–4 December 2018, p. 8.

technological environments, as well as accessible and understandable data-processing methods subject to external scrutiny. At the same time, particular emphasis is placed on fairness and impartiality, ensuring that AI systems do not compromise the integrity of judicial decision-making. Finally, the principle of human control underscores that AI systems must remain subject to meaningful human oversight. Judicial professionals must retain the ability to review AI-assisted decisions and the underlying data, without being bound by such outputs. Users must be clearly informed whether AI-generated solutions are binding, of their right to legal advice and access to a court, and of any use of AI in their case, with the possibility to object and request direct judicial review in accordance with Art. 6 of the ECHR.³⁰⁸ In this respect, the Charter reinforces the primacy of human rights by ensuring that AI remains a supportive tool rather than a substitute for judicial authority.

While the European ethical Charter provided an ethics-based framework specific to the judicial context, subsequent efforts of the COE moved towards a broader approach to the protection of human rights and fundamental freedoms in the use of algorithmic systems. An important step in this regard is *Recommendation CM/Rec(2020)1 of the Committee of Ministers*, which calls on Member States to review their legislative and policy frameworks, as well as their practices in the procurement, design, development and deployment of algorithmic systems, and to ensure their alignment with the accompanying guidelines. On this basis, States are encouraged to establish regulatory frameworks ensuring that private sector actors respect human rights in accordance with international standards, and to equip supervisory authorities with adequate resources and powers. The Guidelines provide a practical framework structured around six key areas—ranging from institutional arrangements to transparency, accountability and research—through which Member States and both public and private actors can develop a human rights-based and trustworthy algorithmic environment. Particular attention is given to the obligations of both State and non-State actors in response to the risks associated with the design, development and use of AI systems. These obligations primarily concern the establishment of effective mechanisms for prevention, monitoring, accountability and access to remedies in cases of fundamental rights violations.³⁰⁹ In this context, the Guidelines em-

³⁰⁸ *Ibid*, pp. 9–12.

³⁰⁹ Council of Europe – Committee of Ministers: *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, 8 April 2020.

phaise the importance of systematic human rights impact assessments and transparent documentation throughout the entire lifecycle of AI systems.

An important complement to this normative line is the Commissioner for Human Rights' recommendation "*Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights*." The recommendation promotes a human rights-based approach to AI, aimed at ensuring its beneficial use while mitigating associated risks. It prioritises the protection of human rights over competing interests and outlines ten key steps to guide Member States in regulating AI. In particular, it emphasises the need for human rights impact assessments, public participation, transparency and disclosure of information on AI systems. It further highlights the importance of non-discrimination, equality, data protection and privacy, as well as the safeguarding of fundamental freedoms, including freedom of expression, assembly, association and the right to work. The recommendation also calls for independent oversight mechanisms and access to effective remedies in cases of rights violations, while underlining the responsibility of States to ensure compliance by private actors and to promote AI literacy.³¹⁰

The above-mentioned instruments illustrate the COE's gradual shift from ethics- and principle-based approaches towards the development of more structured regulatory frameworks for AI governance. At the same time, they contributed to consolidating key normative principles, including human dignity, transparency, accountability, non-discrimination, human oversight, and access to effective remedies. Nevertheless, these instruments remain predominantly soft law in nature and provide only limited guidance regarding enforcement, institutional supervision, liability, or the regulation of rapidly evolving AI systems.

An important step towards the development of binding AI regulation within the COE was *Resolution No. 2341 (2020) of the Parliamentary Assembly on the need for democratic governance of artificial intelligence*. The Resolution recognised that voluntary ethical principles and self-regulation are insufficient to ensure effective protection of human rights in the context of AI. It therefore called for the development of a cross-sectoral regulatory framework grounded in the protection of human rights, democracy, and the rule

³¹⁰ Council of Europe – Commissioner for Human Rights: *Unboxing Artificial Intelligence: 10 steps to protect Human Rights*, May 2019, pp. 7–15. Online: <https://rm.COE.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64>.

of law.³¹¹ In this regard, the Resolution emphasised the need to strengthen safeguards aimed at preventing human rights violations arising from the use of AI systems, particularly through independent oversight mechanisms.³¹²

These developments ultimately led to the establishment of two key committees: the Ad Hoc Committee on Artificial Intelligence (hereinafter “CAHAI”) and the Committee on Artificial Intelligence (hereinafter “CAI”), whose work culminated in the adoption of the Framework Convention, incorporating many of the values and principles articulated in the preceding instruments. While the CAHAI carried out the analytical and consultative phase, including the assessment of feasibility, mapping of existing norms and practices, and consideration of possible regulatory approaches,³¹³ the CAI subsequently assumed the normative role and negotiated the binding text of the Framework Convention.³¹⁴ Both bodies were characterised by the involvement of a broad range of stakeholders. In addition to representatives of the Member States of the COE, the process included several non-European States sharing the values and objectives of the COE, including Argentina, Australia, Canada, Japan, the United States, and Uruguay. Furthermore, observers from international organisations, particularly the EU, OECD and the UN, as well as independent experts, academia, civil society, and private-sector actors participated in the drafting process. This inclusive model made it possible to reflect diverse regulatory traditions, technical considerations, and human rights risks arising across the lifecycle of AI systems.³¹⁵

³¹¹ Council of Europe – Parliamentary Assembly: *Need for democratic governance of artificial intelligence*, Resolution 2341 (2020), 22 October 2020, para. 10–11.

³¹² *Ibid.*, para. 15.

³¹³ Council of Europe – Ad Hoc Committee on Artificial Intelligence: *Feasibility Study*, 17 December 2020, CAHAI(2020)23; Council of Europe – Ad Hoc Committee on Artificial Intelligence: *Possible elements of a legal framework on artificial intelligence, based on the Council of Europe’s standards on human rights, democracy and the rule of law*, 3 December 2021, CAHAI(2021)09rev.

³¹⁴ Council of Europe – Committee on Artificial Intelligence: *Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, 7 July 2023, CAI(2023)18.

³¹⁵ For a more detailed analysis of the preparatory work relating to the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, see BAKOŠOVÁ, L.: *Green Ambitions in the Era of Artificial Intelligence*. Praha: Nakladatelství Leges, 2024, pp. 156–161; MORAWSKA, E. H.: *Council of Europe standards and activities related to AI: towards a Framework Convention on AI and human rights?* In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, pp. 27–43.

The development of the Framework Convention therefore illustrates how soft-law instruments can perform a preparatory and norm-shaping function, gradually consolidating principles that may subsequently be translated into binding legal obligations. In the COE context, this progression is particularly visible in the movement from sector-specific ethical guidance towards a cross-cutting treaty framework grounded explicitly in human rights, democracy and the rule of law.

7.2.2. *The Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*

The *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law* represents the first international treaty comprehensively addressing the regulation of selected aspects of the use of AI systems. It establishes a legal framework governing activities throughout the entire lifecycle of AI systems, reflecting the potential risks associated with their development and deployment while simultaneously creating conditions conducive to responsible innovation. In this respect, the Convention was conceived not merely as a regional instrument of the COE, but as a globally applicable framework based on shared principles and common standards. This intention is expressly reflected in the Preamble, where the Parties declare their “conviction of the need to give priority to establishing a globally applicable legal framework setting out common general principles and rules governing activities within the lifecycle of AI systems, while effectively safeguarding shared values and harnessing the benefits of AI to promote those values in a manner conducive to responsible innovation.”³¹⁶ The Convention thus reflects a broader willingness of States to address AI governance through a supranational human rights framework rather than relying exclusively on fragmented domestic approaches.³¹⁷ Although the Convention represents an important milestone in the international governance of AI, it has not yet entered into force. As of the end of April 2026, it had been signed by eighteen States³¹⁸ and the EU, but the minimum number of ratifications required un-

³¹⁶ *Framework Convention*, Preamble.

³¹⁷ NARDOCCI, C.: *Artificial Intelligence at the Crossroads between the European Union & the Council of Europe: Who Safeguards What & How?* In: *Italian Journal of Public Law*, Vol. 16, Issue 1, 2024, p. 186.

³¹⁸ E.g. Andorra, Canada, Georgia, Iceland, Israel, Japan, Liechtenstein, Moldova, Montenegro, Norway, Switzerland, the United Kingdom of Great Britain and Northern Ireland, or the United States of America.

der the Convention, including ratification by at least three member States of the COE, has not yet been reached.

The objective of the Convention is to ensure that activities within the lifecycle of AI systems remain fully consistent with human rights, democracy, and the rule of law. To achieve these objectives, the Parties undertake to adopt or maintain appropriate legislative, administrative, or other measures proportionate to the severity and likelihood of adverse impacts on human rights, democracy, and the rule of law.³¹⁹ As explained in the *Explanatory Report*, the Convention does not seek to regulate all aspects of activities related to the lifecycle of AI systems, nor does it aim to regulate AI technologies as such. Likewise, the Framework Convention does not address the economic or market-related dimensions of AI systems.³²⁰ Instead, it expressly recognises its framework character, allowing for the future adoption of additional instruments addressing specific issues connected with activities throughout the lifecycle of AI systems.

The Framework Convention applies to public authorities and to private actors acting on their behalf, but it does not automatically extend to the activities of private actors as such. With regard to other private entities, the Convention grants States considerable flexibility, allowing them to determine, through a declaration, how risks and impacts arising from activities within the lifecycle of AI systems should be addressed. In particular, States may choose either to apply the principles and obligations set out in the Framework Convention or to adopt other appropriate measures.³²¹ However, neither the Convention nor the Explanatory Report clarifies what such “other appropriate measures” may entail or the precise scope of this obligation.³²²

This basic delineation of scope is further complemented by a number of explicit exclusions defining situations in which the Convention does not apply. Activities related to national security are excluded, provided that they remain consistent with international law and democratic principles. Matters

³¹⁹ *Framework Convention*, Art. 1

³²⁰ Council of Europe – Committee on Artificial Intelligence: *Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, Vilnius, 5 September 2024, Council of Europe Treaty Series – No. 225 (hereinafter “*Explanatory Report to the Framework Convention*”), para. 5 a 12.

³²¹ *Framework Convention*, Art. 3 (1) (b).

³²² BABICKÁ, K., GIACOMIN, C.: *Understanding the Scope of the Council of Europe Framework Convention on AI*. *Opinio Juris*, 5 November 2024. Online: <https://opiniojuris.org/2024/11/05/understanding-the-scope-of-the-council-of-europe-framework-convention-on-ai/?utm>.

concerning national defence are likewise exempted. In addition, the Convention does not apply to AI research and development activities unless such activities involve testing or conduct capable of interfering with human rights, democracy, or the rule of law.³²³

The fundamental principles and obligations of the Parties constitute a central element of the Framework Convention. In accordance with Arts. 4 and 5, Parties are required to adopt or maintain measures ensuring that activities within the lifecycle of AI systems remain consistent with their international and domestic obligations relating to the protection of human rights and democratic processes. In the context of international human rights protection, the Framework Convention builds upon the legal frameworks established by major international human rights instruments adopted at both the universal and regional levels.³²⁴ It does not seek to create new human rights or alter the scope of existing guarantees. Rather, its objective is to strengthen the implementation of existing obligations through the legally binding provisions contained in Chapters II to VI, thereby responding to the new challenges arising from the use of AI systems. At the same time, the Convention may contribute to the progressive specification of existing human rights in AI-related contexts, particularly by strengthening requirements concerning transparency, procedural safeguards, privacy and protection against algorithmic discrimination. Rather than creating new rights as such, these developments illustrate how existing guarantees may acquire more specific dimensions in response to technological change.

The abovementioned general obligations are further complemented by a set of specifically articulated principles, including human dignity and individual autonomy, transparency and oversight, accountability, equality and non-discrimination, reliability, protection of privacy and personal data, as well as the promotion of safe innovation, which together shape the normative structure of the Convention. The principles enshrined in Arts. 7 to 13 constitute a comprehensive ethical and legal framework linking the technical and value-based dimensions of AI regulation. Their common denominator is the emphasis placed on the anthropocentric orientation of technological development, namely the preservation of human primacy over the systems created by humans themselves. These principles do not merely represent a set of abstract values, but establish a practical normative compass intended to guide both States and private actors in the development and use of AI sys-

³²³ *Framework Convention*, Art. 3.

³²⁴ See *Chapter 3. Normative Framework of International Human Rights Law*.

tems. Their effective implementation is essential to ensuring that technological progress remains balanced with the protection of human rights, democracy, and the rule of law.

In this context, reference should also be made to the *COE Handbook on Human Rights and Artificial Intelligence*, which provides practical guidance on the application of existing human rights standards to the development and use of AI systems. Rather than providing an exhaustive legal analysis, the Handbook offers practical guidance on the application of the ECHR, the European Social Charter, and related human rights standards throughout the lifecycle of AI systems, particularly in the context of public governance. It introduces key technical concepts, outlines relevant human rights principles, and proposes a structured framework for assessing the human rights implications of AI systems across different sectors, including through the identification of risks, applicable legal standards, and emerging good practices. Adopted by the Steering Committee for Human Rights at its plenary meeting in November 2025, the Handbook complements and operationalises the principles reflected in the Framework Convention, thereby reinforcing the coherence of the COE's emerging AI governance architecture.³²⁵

In accordance with Chapter IV, each Party is required, to the extent mandated by its international obligations and in accordance with its domestic legal system, to adopt or maintain measures enabling affected persons to seek effective remedies. Such measures include, in particular, the documentation and provision of relevant information concerning AI systems capable of significantly affecting the exercise of human rights, both to competent authorities and, where appropriate, to affected individuals. The information provided must be sufficiently specific to enable individuals to challenge decisions made, or substantially influenced, by AI systems, or to contest the use of such systems itself.³²⁶ These disclosures must be comprehensible, context-appropriate, and capable of enabling individuals to effectively exercise their rights in proceedings affecting them. At the same time, exceptions to transparency obligations may be established where required for reasons of public order, security, or other important public interests. The Convention further requires that affected persons have access to the possibility of submitting complaints

³²⁵ Council of Europe – Steering Committee for Human Rights: *CDDH Handbook on Human Rights and Artificial Intelligence*, CDDH(2025)R103 Addendum 3, 12 January 2026, para. 4–5. Online: <https://rm.coe.int/steering-committee-for-human-rights-cddh-cddh-handbook-on-human-rights/48802a2f59>.

³²⁶ *Framework Convention*, Art. 14.

to competent authorities. However, the precise approach to be taken to answer the question of whether the rights of applicants have been violated is left outside the scope of this international treaty, leaving room for a variety of models. Thus, the Convention merely establishes a general procedural vector by guaranteeing the availability of a remedy.³²⁷ At the same time, the CAI emphasised in its report that obligations relating to effective remedies and impact assessments do not automatically apply to all AI systems. Their application depends on whether a given AI system is capable of significantly affecting the exercise or enjoyment of human rights. The assessment of whether this threshold is met remains within the discretion of the Parties, which must take into account their existing legal frameworks and the specific circumstances of the particular context.³²⁸

With regard to procedural safeguards, the Parties are required to ensure that, in situations where an AI system significantly affects the exercise of human rights, individuals have access to effective legal remedies and procedural guarantees in accordance with international and domestic law.³²⁹ The Parties are therefore obliged to preserve and adapt existing procedural mechanisms, particularly the possibility of human review of decisions that are “significantly affected” by the use of AI systems. In this way, the Convention seeks to prevent fully automated decision-making without the possibility of human intervention and to ensure accountability and reviewability of decision-making processes. The second part of the provision extends procedural safeguards to the sphere of human interaction with AI systems. The obligation to inform users that they are interacting with a machine is of particular importance for the protection of autonomy, informed consent, and trust. In this respect, Art. 15 reinforces the requirement of transparent and controllable interaction between humans and technology, thereby building upon the broader principles of accountability, oversight, and transparency.³³⁰ These provisions strengthen the responsibility of actors involved throughout the AI lifecycle and constitute an essential procedural safeguard protecting individuals against opaque or disproportionate uses of AI in decision-making processes.

³²⁷ MARTYNOVA, E.: *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary*. In: BELLI, L., GASPAR, W. (eds.): *AI from the Global Majority: Official Outcome of the UN IGF Data and Artificial Intelligence Governance Coalition*. Rio de Janeiro: FGV Direito Rio, 2024, p. 5.

³²⁸ *Explanatory Report to the Framework Convention*, para. 99–101.

³²⁹ *Framework Convention*, Art. 15.

³³⁰ *Explanatory Report to the Framework Convention*, para. 103–104.

Effective risk assessment and the subsequent management of impacts constitute an integral part of the responsible use of AI. The Framework Convention establishes a preventive mechanism aimed at the early identification and mitigation of risks that may arise in the design, development, and use of AI systems. The deployment and use of such systems should therefore be based on human rights risk assessments and guided by the principles of transparency and accountability.³³¹ The objective of Art. 16 is to ensure a coherent approach to the identification, analysis, and evaluation of these risks, as well as to the assessment of the impacts of AI systems throughout their entire lifecycle. At the same time, the provision is based on the assumption that individual States are best placed to adopt relevant regulatory decisions, given their particular knowledge of their own legal, political, economic, social, cultural, and technological contexts. For this reason, States should retain an appropriate degree of flexibility in the management and regulation of processes related to the use of AI systems, enabling them to take into account their specific legal, societal, and technological circumstances.³³²

Building upon the above-mentioned obligations, the Parties are required to establish legal and institutional frameworks enabling the continuous monitoring of risks and the adoption of measures aimed at their mitigation. Such measures must be proportionate and context-sensitive, taking into account the intended use of AI systems, the severity and likelihood of potential impacts, as well as the perspectives of persons whose rights may be affected. Risk management is conceived as an iterative process operating throughout the entire lifecycle of AI systems, accompanied by ongoing risk monitoring, documentation, and, where appropriate, testing of systems both prior to their initial deployment and following substantial modifications.³³³ As *Chang* observes, this represents a particularly notable feature of the Framework Convention's approach, which places emphasis on the involvement of relevant stakeholders, especially individuals whose rights may be affected.³³⁴

³³¹ BENKO, R.: *The Impact of Artificial Intelligence on Fundamental Rights and Freedoms and Vice-Versa*. In: KLUČKA, J., BAKOŠOVÁ, L. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Nakladatelství Leges, 2021, p. 49.

³³² *Explanatory Report to the Framework Convention*, para. 106.

³³³ *Framework Convention*, Art. 16 (2) (a) – (g).

³³⁴ CHANG, CH.: *The First Global AI Treaty: Analyzing the Framework Convention on Artificial Intelligence and the EU AI Act*. In: *University of Illinois Law Review Online*, Vol. 86, 2024, p. 95.

Following the identification of risks, the Parties are required to adopt or maintain mechanisms for the management of impacts and risks associated with AI systems, including so-called risk and impact assessments of AI. These mechanisms encompass the adoption of regulatory, technical, organisational, and educational measures aimed at reducing the likelihood of harm, as well as procedures for evaluating the effectiveness of the measures adopted. In this context, particular attention should be given to the *Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law* (hereinafter “HUDERIA”), adopted by the CAI in November 2024. HUDERIA provides States and private actors with a structured approach to assessing the risks and impacts of AI systems, specifically tailored to the protection and promotion of human rights, democracy, and the rule of law. The methodology originated within the work of the CAHAI, particularly its Policy Development Group, which mandated the Alan Turing Institute to prepare an initial draft operationalising a Human Rights, Democracy and Rule of Law Impact Assessment Model. From a legal perspective, HUDERIA constitutes a non-binding guidance instrument and is not intended to serve as an interpretative tool for the Framework Convention. Nevertheless, Parties may use or adapt the methodology, in whole or in part, when developing new approaches to AI risk assessment or modifying existing ones in accordance with their domestic legal frameworks, provided that they fully comply with their obligations under the Framework Convention, particularly with regard to the minimum requirements relating to risk and impact management.³³⁵

HUDERIA is based on a socio-technical approach linking the technical characteristics of AI systems with the broader social, legal, and political context of their development and use, while assessing factors such as severity, scale, reversibility, and likelihood of potential impacts.

The methodology is structured around four key components, the first of which is the Context-Based Risk Analysis (COBRA). This initial phase focuses on the systematic identification of risk factors associated with AI systems and their potential impact on human rights, democracy, and the rule of law. COBRA examines the intended purpose of the AI system, its technical design, deployment context, degree of human oversight, and the na-

³³⁵ Council of Europe - Committee on Artificial Intelligence: *Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law (Huderia Methodology)*, CAI(2024)16rev2, 28 November 2024, p. 3.

ture of the data processed. It also identifies potentially affected individuals or groups and assesses the severity and likelihood of potential harms in order to determine the appropriate scope of further risk assessment and whether the deployment of the system is justified in light of its potential risks and benefits.³³⁶

The second component, the Stakeholder Engagement Process, complements the risk assessment by incorporating the perspectives of individuals and groups potentially affected by AI systems, including vulnerable groups. It focuses on identifying affected stakeholders, addressing power imbalances between developers and users, and ensuring meaningful, transparent, and non-discriminatory participation throughout the assessment process. In this way, Stakeholder Engagement Process strengthens both the quality and legitimacy of risk and impact assessments.³³⁷

The third component, the Risk and Impact Assessment, provides a more detailed evaluation of the potential and actual impacts of AI systems on human rights, democracy, and the rule of law, particularly in relation to systems identified as high-risk during the COBRA phase. The assessment analyses how harms may arise in a specific context and evaluates factors such as severity, scale, reversibility, and likelihood in order to prioritise risks and prepare appropriate mitigation measures.³³⁸

Finally, the Mitigation Plan establishes concrete measures aimed at preventing and addressing adverse impacts of AI systems. It prioritises actions according to the seriousness and probability of harm and follows a hierarchy focused on prevention, mitigation, restoration, and compensation. The plan also emphasises transparency, accountability, access to remedies, and the continuous monitoring and reassessment of mitigation measures throughout the lifecycle of the AI system.³³⁹

The entire risk and impact assessment process is conceived as iterative in nature. The methodology expressly recognises that changes in data, algorithms, deployment methods, or the broader legal and social environment may substantially alter the risk profile of an AI system. For this reason, it requires continuous “Iterative Review”, involving the repeated reassessment of risks, the updating of mitigation measures, and, where necessary, the suspension or substantial revision of the use of the AI system. This phase is intend-

³³⁶ *Ibid*, pp. 7–12.

³³⁷ *Ibid*, pp. 13–15.

³³⁸ *Ibid*, pp. 16–18.

³³⁹ *Ibid*, pp. 19–21.

ed to remain flexible and responsive to both developmental and operational conditions, supported by ongoing monitoring to ensure that the protection of human rights, democracy, and the rule of law is maintained throughout the entire lifecycle of the system, from design to decommissioning.³⁴⁰ In this respect, HUDERIA represents an important normative bridge between the abstract principles of the Framework Convention and the practical processes of AI risk governance. However, its implementation model allows for adaptation to different contexts, leaving considerable discretion in the concrete implementation of impact assessments and therefore representing a flexible and adaptable tool.³⁴¹

Where the use of certain AI systems is considered incompatible with the protection of human rights, the functioning of democracy, or the principles of the rule of law, a Party may introduce a moratorium, a complete prohibition, or other appropriate measures. The assessment of such incompatibility remains entirely within the discretion of the Party concerned.³⁴² In making this determination, Parties should carefully consider whether effective risk mitigation measures are available and, where this is not the case, adopt stricter regulatory responses. Such measures should be accompanied by appropriately designed review mechanisms enabling their regular reassessment, modification, or withdrawal, for example where risks diminish or new unacceptable practices emerge. In such situations, public consultations are of particular importance in ensuring transparency, legitimacy, and public acceptance.³⁴³ Effective risk assessment and impact management therefore constitute a fundamental prerequisite for the practical implementation of the principles of the Framework Convention. Their effective operation requires the establishment of clearly defined institutional and oversight mechanisms capable of ensuring that the principles and measures adopted are translated into practical reality.

The practical effectiveness of the Framework Convention is closely linked not only to the substantive standards it establishes, but also to the institutional and procedural mechanisms designed to secure their implementation and

³⁴⁰ *Ibid*, pp. 22–23.

³⁴¹ BALDINELLI, L.: *Artificial Intelligence in Public Administration: The Importance of AI Impact Assessments between the Framework Convention and the AI Act*. In: Italian Journal of Public Law, Vol. 17, No. 4, 2025, p. 1381; PUSZTAHELYI, R.: *Some Reflections on Fundamental Rights Impact Assessment for High-Risk AI Systems (FRIA)*. In: The Juridical Current Journal, Vol. 103, No. 4, 2025, pp. 24–25.

³⁴² *Framework Convention*, Art. 16 (4).

³⁴³ *Explanatory Report to the Framework Convention*, para. 112.

supervision in practice. Under Chapter VI, the States Parties are required to ensure the implementation of the Convention without discrimination of any kind and in accordance with their international human rights obligations. The Convention places particular emphasis on the need to consider the specific needs and vulnerabilities of persons with disabilities and children, while also recognising the importance of public discussion and inclusive consultation in addressing significant issues connected with the use of AI systems. Such discussions are expected to involve a broad range of stakeholders, including the general public, industry experts, academia, national human rights institutions, and civil society organisations. These participatory processes play an important role in ensuring that AI systems remain consistent with universal human rights standards and are responsive to concerns relating to human rights, democracy, and the rule of law. By incorporating diverse perspectives, they further contribute to the development of related policies and regulatory initiatives.³⁴⁴

To support the effective implementation of the Framework Convention, States Parties are also encouraged to promote digital literacy and specialised digital skills among both the general public and professionals involved in identifying and managing risks associated with AI systems.³⁴⁵ This provision reflects broader developments in both binding and non-binding instruments governing digital technologies, including AI, such as similar commitments contained in the Global Digital Compact and the Artificial Intelligence Act.

Chapter VI further expressly provides that no provision of the Convention may be interpreted as limiting or derogating from existing human rights protections, nor as preventing States from adopting broader safeguards than those established under the Convention itself.³⁴⁶

This socially and value-oriented implementation framework is complemented by procedural and institutional mechanisms aimed at monitoring compliance with the Convention and fostering international cooperation among the Parties. For this purpose, the Convention establishes a Conference of the Parties tasked with facilitating its effective implementation and application. The implementation and supervisory mechanisms established under the Convention are examined in greater detail in *Chapter 9.2.1. Governance and Oversight under the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*.

³⁴⁴ *Explanatory Report to the Framework Convention*, point 120.

³⁴⁵ *Framework Convention*, Art. 20.

³⁴⁶ *Ibid*, Art. 21–22.

The Convention also addresses its relationship with EU law. Under Art. 27 (2), Parties that are members of the EU may apply EU rules governing matters within the scope of the Convention in their mutual relations, provided that this does not prejudice the object and purpose of the Convention or its application in relations with other Parties. The same applies to non-EU Parties to the extent that they are bound by relevant EU rules.³⁴⁷ This provision reflects the advanced level of legal integration within the EU, particularly in the field of AI regulation, while also illustrating the growing interaction between the COE framework on AI and the EU's regulatory approach, especially the AI Act. Although the Convention seeks to ensure coherence with the EU legal framework, it nevertheless remains a distinct international instrument adopted within the institutional structure of the COE, whose membership and normative mandate differ from those of the EU.³⁴⁸

Overall, the Framework Convention confirms rather than displaces the continuing relevance of existing international human rights law in the governance of AI. Its principal contribution lies in translating established human rights obligations into more specific procedural and governance requirements concerning transparency, remedies, risk and impact assessment, human oversight and accountability. At the same time, the considerable discretion left to Parties, particularly regarding private-sector activities and the implementation of risk-management obligations, may result in uneven levels of protection. The effectiveness of the Convention will therefore depend not only on its entry into force and subsequent ratification, but also on the manner in which its general obligations are implemented within domestic legal systems.

³⁴⁷ *Explanatory Report to the Framework Convention*, para. 147–148.

³⁴⁸ ROTENBERG, M.: *Introductory Note to Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. In: *International Legal Materials*, Vol. 64, 2025, p. 861.

III. INSTITUTIONAL FRAMEWORKS FOR THE PROTECTION OF HUMAN RIGHTS IN THE AGE OF ARTIFICIAL INTELLIGENCE

Beyond the development of regulatory frameworks, the effective protection of human rights in the age of AI also depends on the functioning of institutional and procedural mechanisms responsible for monitoring, interpreting, and enforcing human rights standards.³⁴⁹ While IHRL provides a normative foundation for addressing the challenges posed by AI, its practical impact is largely determined by the capacity of existing institutions to apply these norms in technologically complex and rapidly evolving contexts. In this regard, both universal and regional human rights mechanisms play a crucial role. However, their mandates were not originally designed to address the specific risks associated with AI systems. At the same time, recent developments point towards the emergence of new advisory and cooperative mechanisms aimed at bridging existing gaps and strengthening the institutional framework for AI governance.

This Part therefore examines the role of existing human rights institutions in responding to the challenges posed by AI and considers emerging institutional responses and future perspectives for ensuring effective human rights protection in the digital age, with particular attention devoted to the practice of key judicial and quasi-judicial bodies at both the universal and regional levels.

³⁴⁹ For more details see *Chapter 4 – Institutional Framework for the Protection of Human Rights*.

8. THE UNITED NATIONS HUMAN RIGHTS SYSTEM AND ARTIFICIAL INTELLIGENCE

The UN human rights system constitutes the principal universal framework for the promotion, monitoring and protection of human rights. In the context of AI, this system has not been fundamentally redesigned; rather, it has been progressively adapted through the interpretative practice of its existing mechanisms. Through general comments and individual communications, treaty bodies have progressively elaborated standards concerning privacy, non-discrimination, procedural fairness and the protection of vulnerable groups in the digital environment, which were already discussed in the previous Part.

Nevertheless, the increasing relevance of AI within the field of human rights also places growing pressure on the existing supervisory architecture of the UN treaty body system. Although treaty bodies have demonstrated a considerable capacity to adapt human rights norms through interpretation, their procedures and working methods were developed in a legal and institutional environment preceding the emergence of large-scale digital technologies and algorithmic governance. In practice, this creates difficulties in addressing technologically complex forms of interference with human rights, particularly where such impacts are diffuse, cross-border and closely linked to the activities of private actors. The UN treaty bodies have not yet produced a substantial body of views directly addressing the specific human rights challenges arising from AI systems as such. At the same time, these limitations should not be understood solely as institutional shortcomings.

8.1. The Emergence of AI-Specific Mechanisms within the United Nations

The limitations of existing human rights mechanisms in addressing AI-related challenges are increasingly reflected in contemporary debates within the UN system. In this context, the Global Digital Compact represents an important step in the gradual development of global AI governance. Although non-binding in nature, the Compact emphasises the need for stronger in-

ternational cooperation and more coherent responses to the human rights implications of digital technologies. While recognising the important role of the Office of the High Commissioner for Human Rights in providing expertise on human rights and technology,³⁵⁰ the Compact also highlights the absence of a comprehensive framework capable of responding to the cross-border and systemic risks associated with AI systems.³⁵¹ These developments have contributed to growing calls for the establishment of specialised AI-related mechanisms within the UN framework, including proposals for a UN Special Rapporteur on Artificial Intelligence and Human Rights or a dedicated coordination office for AI governance.

One of the most notable recent developments within the UN framework is the creation of the Independent International Scientific Panel on Artificial Intelligence. The establishment of the Panel reflects broader efforts to strengthen institutional capacity and expert coordination in the field of AI governance at the global level. Composed of forty independent experts representing different regions and disciplines, the Panel is tasked with preparing annual evidence-based reports examining the opportunities, risks and societal impacts associated with AI technologies, supplemented where appropriate by thematic analytical briefs. The institutional design of the Panel emphasises scientific independence, multidisciplinary and geographical inclusiveness. Its structure combines co-chairs from developed and developing states, vice-chairs and ad hoc working groups, while also allowing cooperation with external experts and stakeholders. Rather than exercising regulatory or supervisory functions, the Panel primarily operates as an advisory and knowledge-based mechanism intended to facilitate informed international dialogue and support policy coordination within the broader framework of the Global Dialogue on AI Governance.³⁵² At the same time, the mandate of the Panel remains limited from the perspective of direct human rights protection. Its role is centred on expertise, assessment and policy guidance rather than the adoption of binding standards or the provision of oversight and

³⁵⁰ UN General Assembly: Pact for the Future – Annex – Global Digital Compact, A/RES/79/1, 22, September 2024, para. 24.

³⁵¹ GOEL, B., NILKANTH, D., SINGH, S.: *AI Governance Beyond 2025: UN Pathways and Implications*. Geneva: Geneva Graduate Institute, 2025, p. 30. Online: https://www.graduateinstitute.ch/sites/internet/files/2025-09/ARP35_Report--2---Shreya-Singh.pdf.

³⁵² UN General Assembly: *Terms of reference and modalities for the establishment and functioning of the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance*, A/RES/79/325, 26 August 2025, para. 1.

remedies in cases of human rights violations. Consequently, although the Panel may contribute to a more informed and coordinated global approach to AI governance, questions remain regarding its capacity to address issues of accountability and effective enforcement, which continue to represent central challenges in the protection of human rights in the age of AI.

Complementing the work of the Scientific Panel, the proposed governance architecture further includes the Global Dialogue on AI Governance, established within the framework of the GDC. The Dialogue is envisaged as an annual multistakeholder forum, expected to convene from 2026 onwards in connection with the AI for Good Summit. Its primary objective is to provide a platform for cooperation and exchange among States, international organisations, private actors, academia and civil society concerning the opportunities, risks and regulatory implications associated with AI technologies. In this respect, the Dialogue is intended to support more inclusive discussions on global AI governance and to promote the use of AI in advancing the Sustainable Development Goals.³⁵³ From an institutional perspective, the Dialogue reflects an increasing emphasis on coordination, participation and policy exchange within the emerging framework of global AI governance. At the same time, its role remains predominantly consultative and discussion oriented. Unlike judicial or supervisory mechanisms within the field of human rights protection, the Dialogue does not possess monitoring, enforcement or complaint-handling powers. Its contribution therefore lies primarily in facilitating international engagement and policy coordination rather than in ensuring direct protection against AI-related human rights violations.

Collectively, these initiatives reflect the gradual emergence of a more coordinated and institutionalised approach to AI governance within the UN framework. Nevertheless, their primary focus remains on expert cooperation, policy coordination and international dialogue rather than on the development of mechanisms specifically aimed at safeguarding human rights in the context of AI. As a result, increasing attention has been devoted to proposals advocating the creation of more specialised institutional structures within the UN system capable of addressing the human rights implications of AI in a more direct and effective manner.

Alongside the emergence of new coordination and advisory structures, proposals have also appeared for mechanisms more directly focused on the human rights implications of AI. In this regard, the Center for AI & Digi-

³⁵³ *Ibid*, para. 4.

tal Policy has called for the establishment of a Special Rapporteur on Artificial Intelligence and Human Rights within the framework of the UN Human Rights Council.³⁵⁴ Such a mandate would operate within the system of special procedures, one of the principal monitoring mechanisms of the UN human rights framework. Special Rapporteurs and other mandate holders are independent experts entrusted with examining thematic or country-specific human rights issues through activities such as country visits, communications with states, thematic reporting and expert consultations. Supported by the OHCHR, they also contribute to the interpretation and progressive development of international human rights standards and often play an important role in drawing international attention to emerging issues.³⁵⁵ In the context of AI, a dedicated mandate could contribute to more systematic monitoring of AI-related human rights risks and strengthen international discussions concerning the governance of emerging technologies. However, the impact of such a mechanism would remain influenced by the broader limitations characteristic of the special procedures system, particularly its non-binding and primarily recommendatory nature. As a result, although a Special Rapporteur on AI and Human Rights could reinforce visibility, coordination and normative guidance within the UN framework, it would not in itself resolve broader questions relating to accountability, enforcement and effective remedies for AI-related human rights violations.

A further proposal emerging within the UN framework concerns the establishment of a dedicated AI Office with a system-wide coordination role in the field of digital and AI governance. The idea was advanced by the High-Level Advisory Body on Artificial Intelligence in its final report³⁵⁶ and subsequently reflected in discussions surrounding the Global Digital Compact.³⁵⁷

³⁵⁴ Center for AI & Digital Policy: *Statement of the Center for AI & Digital Policy to the Vice-President of the United Nations Human Rights Council To support the Establishment of the Special Rapporteur for AI and Human Rights*, November 2025. Online: <https://www.caidp.org/app/download/8585853363/CAIDP-UNHRC-VPRusu-AI-Treaty-09112025.pdf?t=1763220293&utm>.

³⁵⁵ UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007, para. 39–54; UN Human Rights Council: *Manual of Operations*, August 2008. Online: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.ohchr.org%2Fsites%2Fdefault%2Ffiles%2FDocuments%2FHRCBodies%2FSP%2FManual_August_2008.doc&wdOrigin=BROWSELINK.

³⁵⁶ United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, pp. 70–72.

³⁵⁷ UN General Assembly: *Pact for the Future – Annex – Global Digital Compact*, A/RES/79/1, 22, September 2024, para. 72.

According to these proposals, the Office would be situated within the UN Secretariat and operate under the authority of the UN Secretary-General as a flexible coordination mechanism rather than as an autonomous regulatory institution. The proposed Office is envisaged primarily as a central coordinating platform linking the different components of the emerging UN AI governance architecture, including the Independent International Scientific Panel on Artificial Intelligence, the Global Dialogue on AI Governance and related initiatives concerning standards exchange, policy coordination and capacity-building. In this respect, its role would focus on facilitating cooperation and information-sharing both within the UN system and among external stakeholders, including states, international organisations, private actors, academia and civil society. The Office would therefore seek to strengthen coherence and coordination in global AI governance rather than exercise direct supervisory or enforcement functions.³⁵⁸

From the perspective of human rights protection, the establishment of such an Office could contribute to greater visibility of AI-related risks and improve institutional coordination concerning the human rights implications of emerging technologies. At the same time, however, the proposed mechanism remains primarily organisational and policy-oriented in nature. It would not possess powers enabling it to determine responsibility for human rights violations, provide remedies to affected individuals or ensure enforcement of international human rights obligations. Consequently, although the creation of an AI Office may represent an important step towards a more coherent institutional approach to AI governance within the UN system, its direct contribution to effective human rights protection would remain inherently limited.

Taken as a whole, recent developments within the UN point to the gradual consolidation of a more articulated, multi-layered institutional landscape for AI governance, within which human rights are increasingly positioned as a central normative reference point. Despite this trajectory, a discernible tension persists between the breadth of normative commitments and the practical capacity of existing institutions. While established human rights mechanisms offer a robust interpretative framework, the structural features of AI systems, including opacity, fragmented chains of responsibility, and cross-border deployment, complicate both the application and the enforcement of these standards. Emerging initiatives have enhanced coordination and fa-

³⁵⁸ United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, pp. 70–72.

cilitated the exchange of expertise and best practices. However, their role remains largely facilitative, functioning primarily as forums for dialogue rather than as bodies endowed with direct supervisory or enforcement powers. Consequently, key issues relating to accountability, effective oversight, and access to remedies continue to lack comprehensive institutional responses, leaving open the question of whether the current trajectory will culminate in a sufficiently effective system of human rights protection in the age of AI.

9. REGIONAL HUMAN RIGHTS MECHANISMS

Valuable insights can also be drawn from regional human rights mechanisms, which tend to respond more flexibly than global institutions and may function as important testing grounds for emerging governance approaches. The following sections examine key regional systems, which increasingly shape both the protection of human rights and the development of AI governance through context-specific and evolving approaches. Before turning to newly established or emerging institutional frameworks, it is important not to overlook the role of regional judicial and quasi-judicial bodies, whose jurisprudence continues to play a crucial role in safeguarding fundamental rights and freedoms. Through their interpretative function, these courts contribute to the dynamic adaptation of existing human rights standards to new technological contexts, including those arising from the development and use of AI.³⁵⁹

9.1. The European Union Institutional Framework for Artificial Intelligence and Fundamental Rights

Within the EU legal order, the protection of fundamental rights is primarily entrusted to the CJEU and the General Court, whose jurisprudence increasingly shapes the relationship between technological regulation, digital governance, and fundamental rights protection. In this regard, the CJEU has progressively developed an important body of case law concerning data protection, surveillance, and automated decision-making, which is highly relevant in the context of AI. This is particularly evident in judgments such as *Digital Rights Ireland*, *Tele2 Sverige*, *La Quadrature du Net*, and *Schrems I and II*, in which the Court established strict limitations on surveillance practices,

³⁵⁹ See, e.g. *Big Brother Watch and Others v. the United Kingdom*, Applications nos. 58170/13, 62322/14 and 24960/15, Judgment of 25 May 2021; *S. and Marper v. the United Kingdom*, Application nos. 30562/04 and 30566/04, Judgment of 4 December 2008; *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland*, Application no. 931/13, Judgment of 27 June 2017; *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023; *Digital Rights Ireland Ltd v Minister for Communications*, Joined Cases C-293/12 and C-594/12, Judgment of 8 April 2014; *Ligue des droits humains v Council of Ministers*, Case C-817/19, Judgment of 21 June 2022.

reinforced the principle of proportionality, and emphasised the necessity of effective judicial protection in data-driven environments.

In *Digital Rights Ireland*, the Court invalidated the Data Retention Directive on the grounds that the general and indiscriminate retention of communications data constituted a particularly serious interference with the rights to privacy and data protection under Arts. 7 and 8 Charter of Fundamental Rights of the EU. The Court stressed that any interference of such magnitude must be governed by clear and precise rules and accompanied by sufficient safeguards capable of preventing abuse and arbitrary access to personal data.³⁶⁰ Building on this reasoning, the Court in *Tele2 Sverige* further clarified that national legislation providing for the general and indiscriminate retention of traffic and location data exceeds the limits of what is strictly necessary in a democratic society. The judgment emphasised the importance of targeted retention measures, independent prior review, and effective oversight mechanisms.³⁶¹

Similar concerns were reiterated in *La Quadrature du Net*, where the Court acknowledged that national security considerations may justify certain surveillance measures, but only under exceptional circumstances and subject to strict proportionality requirements, effective safeguards, and independent supervision.³⁶² In parallel, the *Schrems I* and *Schrems II* judgments significantly strengthened the EU approach to data protection in the context of international data transfers. In these cases, the Court criticised broad and insufficiently limited access powers of public authorities, underlining that effective protection of personal data requires enforceable rights, effective judicial remedies, and protection against disproportionate surveillance practices.³⁶³ Collectively, this jurisprudence establishes core principles highly relevant for AI governance, including the necessity of clear legal frameworks, limitations on indiscriminate data processing and surveillance, independent oversight, proportionality, and robust safeguards against arbitrariness in technologically enabled decision-making systems.

³⁶⁰ *Digital Rights Ireland Ltd v Minister for Communications*, Joined Cases C-293/12 and C-594/12, Judgment of 8 April 2014, para. 65–66.

³⁶¹ *Tele2 Sverige AB Post- och telestyrelsen*, Joined Cases C-203/15 and C-698/15, Judgment of 21 December 2016, para. 134.

³⁶² *La Quadrature du Net and Others*, Joined Cases C-511/18, C-512/18 and C-520/18, Judgment of 6 October 2020, para. 229.

³⁶³ *Schrems*, Case C-362/14, Judgment of 6 October 2015, para. 93–95; *Schrems II*, Case C-311/18, Judgment of 16 July 2020, para. 130–132, 168–177, 179–181, 192–196.

More recently, in *SCHUFA Holding*, the Court adopted a functional interpretation of automated decision-making under EU data protection law, recognising that even probabilistic scoring systems may fall within the scope of Art. 22 GDPR where they effectively determine decisions in practice. Although the case did not directly concern AI systems, it provides an important normative framework for addressing AI-related risks, particularly by exposing the limitations of purely formal human oversight and highlighting the growing regulatory significance of private actors within digital governance structures.³⁶⁴

Currently, the CJEU is deciding the case *Like Company v. Google Ireland* concerning the training of generative AI and the use of chatbots in relation to the press publishers' ancillary copyright, as well as the regulation of text and data mining. In essence, the CJEU is asked to clarify the applicability of the rights of reproduction and making available to the public in connection with the reproduction and display of content in the replies of Google's chatbot where that content is (in part) identical to the content on a newspaper's website. In addition to the AI training process itself, the specific case of the application of Google Gemini and thus the creation of AI output is considered.³⁶⁵ Additionally, the CJEU is currently addressing broader questions concerning automated decision-making, transparency, and judicial oversight in the pending case *Yettel Bulgaria*. The reference raises issues relating to the consumer's right to obtain information about the functioning of AI-driven or automated systems, including the algorithms and parameters underlying automated invoicing, as well as the extent to which courts may require disclosure of source code or "black box" information in judicial proceedings. The case further illustrates the growing interaction between the AI Act, EU consumer protection law, and the fundamental rights guarantees enshrined in Arts. 38 and 47 of the Charter, particularly with regard to transparency, explainability, and effective judicial protection in the context of automated decision-making.³⁶⁶

³⁶⁴ *SCHUFA Holding (Scoring)*, Case C-634/21, Judgment of 7 December 2023, para. 46–60.

³⁶⁵ *Like Company v. Google Ireland*, Case C-250/25, Request for a preliminary ruling, 3 April 2025; HOFFMANN, J.: *Technological Determination of AI-Relevant Press and Copyright Law and Generative Content's Relevance for EU Competition Law – The Referral in Case C-250/25, Like Company v. Google Ireland Ltd.*, Max Planck Institute for Innovation and Competition Research Paper No. 25-19, 2025, p. 3. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5411443.

³⁶⁶ *Yettel Bulgaria*, Case C-806/24, Request for a preliminary ruling, 24 February 2025.

The forthcoming judgments may therefore prove highly significant not only for the interpretation of EU copyright law in the context of generative AI, but also for the broader development of EU approaches to AI governance, particularly with regard to AI training practices, automated decision-making, transparency obligations, and the legal accountability associated with AI-generated outputs and algorithmic systems.

9.1.1. Governance Architecture under the Artificial Intelligence Act

The regulatory approach of the AI Act is complemented by a multi-layered governance structure designed to ensure its consistent implementation and enforcement. It consists of the AI Office, the European Artificial Intelligence Board (hereinafter “EAIB”), the Advisory Forum, the Scientific Panel, and national competent authorities in each Member State.

At the Union level, a central role is entrusted to the AI Office within the European Commission, which is responsible for developing expertise and coordinating the implementation of the Act, particularly in relation to GPAI models.³⁶⁷ According to the Commission Decision establishing the Office, its creation reflects the need to develop Union-level expertise and institutional capacities capable of responding to the rapid evolution and increasing complexity of AI technologies, particularly general-purpose AI models. The AI Office is tasked not only with supporting the implementation and enforcement of the AI Act, but also with monitoring technological and market developments, assessing systemic risks, facilitating international cooperation, coordinating with scientific and regulatory actors, and promoting the development and deployment of trustworthy AI within the Union.³⁶⁸

The Act further establishes the EAIB, composed of representatives of the Member States, with the participation of the AI Office and the European Data Protection Supervisor. The Board functions primarily as a coordination and advisory body aimed at ensuring the consistent application of the AI Act across the Union. Its tasks include facilitating cooperation between national authorities, promoting harmonised administrative practices, supporting the development of guidance and standards, advising the Commission on

³⁶⁷ AI Act, Art. 64 and 53–55.

³⁶⁸ Commission Decision of 24.1.2024 establishing the European Artificial Intelligence Office, Brussels, 24.1.2024, C(2024) 390 final, Arts. 1–7.

emerging technological developments, and contributing to AI literacy and international cooperation.³⁶⁹

To ensure broader stakeholder participation, the AI Act also creates an Advisory Forum composed of representatives of industry, SMEs, civil society, academia, and standardisation bodies. The Forum provides technical expertise and contributes to the development of recommendations and opinions concerning the implementation of the Regulation. In parallel, a Scientific Panel of independent experts supports the Commission and the AI Office in highly technical matters, including the assessment of systemic risks posed by general-purpose AI models, benchmarking, and enforcement-related activities.³⁷⁰

Supranational authorities have an essential role, but the effective implementation and enforcement of this Act frequently require a local presence, placing the responsibility primarily on Member States.³⁷¹ At the national level, Member States are required to designate competent authorities, including market surveillance authorities and notifying authorities, responsible for the supervision and enforcement of the AI Act. These authorities must operate independently and possess adequate technical, financial, and human resources, including expertise in AI technologies, cybersecurity, data protection, and fundamental rights. Member States must also establish single points of contact to facilitate coordination with EU institutions and other national authorities.³⁷²

From a human-rights perspective, this governance architecture reflects a shift towards a hybrid and expertise-driven model, combining legal oversight with technical and scientific input. At the same time, the bodies constituting this governance architecture perform primarily coordination, supervisory, and advisory functions. Their role is therefore best understood as complementary, enhancing the capacity of existing legal frameworks to address the complex and evolving human-rights implications of AI systems. Nevertheless, the governance architecture introduced by the AI Act also rais-

³⁶⁹ AI Act, Art. 66; STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J.: *Artificial Intelligence and International Human Rights Law Developing Standards for a Changing World*, Cheltenham: Edward Elgar Publishing, 2024, p. 64.

³⁷⁰ *Ibid.*, Art. 67–69.

³⁷¹ NOVELLI, C. et al.: *A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities*. In: *European Journal of Risk Regulation*, Vol. 15, 2024, p. 17.

³⁷² AI Act, Art. 70.

es a number of important challenges, particularly with regard to ensuring coherent enforcement across Member States, avoiding regulatory fragmentation, securing sufficient expertise and institutional capacity at the national level, and maintaining an appropriate balance between technological innovation and the protection of fundamental rights. As suggested by *Novelli et al.*, these challenges could be addressed through several institutional reforms, including strengthening the independence of the AI Office, simplifying advisory structures to improve coherence and efficiency, enhancing coordination between EU bodies to avoid overlapping competences, and reinforcing the role of the AI Board in promoting consistent enforcement across Member States. The authors further emphasise the importance of adaptive governance mechanisms capable of continuously monitoring technological developments and disseminating regulatory best practices.³⁷³

9.1.2. Complementary EU Governance Structures Supporting Fundamental Rights Protection

The EU regulatory landscape relevant to AI is not limited to the AI Act itself but also encompasses a broader set of legal instruments governing the digital environment, including the GDPR, the Data Act, the Digital Services Act, and the Digital Markets Act, which collectively contribute to strengthening the protection of fundamental rights in areas affected by automated decision-making, large-scale data processing, and digital platform governance. Together, these instruments form an interconnected regulatory framework aimed at addressing the broader societal and legal implications of digital technologies.

Within this framework, an important role is played by the European Data Protection Board (hereinafter “EDPB”), whose task is to ensure the coherent application of EU data protection rules across the Member States, particularly under the GDPR. Through its interpretative activity, including the adoption of guidelines, opinions, recommendations, and binding decisions in consistency and dispute resolution procedures, the EDPB significantly shapes the practical understanding of rights related to privacy and personal data protection in the digital sphere.³⁷⁴ Its work is especially relevant in areas

³⁷³ NOVELLI, C. et al.: *A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities*. In: *European Journal of Risk Regulation*, Vol. 15, 2024, p. 25.

³⁷⁴ *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on*

involving profiling, automated decision-making, and the processing of large volumes of personal data. Although the EDPB does not function as a judicial or specialised human-rights body and is not empowered to provide individual remedies for human rights violations, its activities have a considerable normative and regulatory impact on the protection of fundamental rights within the EU digital order. In this respect, the Board may be regarded as an important supervisory and interpretative actor contributing to the broader governance framework applicable to AI and digital technologies.

An important institutional actor within the EU framework is the European Data Protection Supervisor (hereinafter “EDPS”), whose mandate has gradually expanded beyond the traditional field of data protection³⁷⁵ in response to the growing regulation of AI systems. Under the AI Act, the EDPS assumes an integrated supervisory role in relation to AI systems used by EU institutions, bodies, offices, and agencies. In this capacity, it may act not only as a supervisory authority, but also as a market surveillance authority and, in certain cases, as a notified body involved in conformity assessment procedures concerning high-risk AI systems. Beyond its direct supervisory responsibilities, the EDPS contributes to the governance framework established under the AI Act through its participation in the activities of the EAIB and by supporting the coherent interpretation and implementation of the regulatory framework through legal and technical expertise.³⁷⁶ Its activities increasingly reflect a preventive and cooperative approach to AI governance based on guidance, institutional support, expert cooperation, and the promotion of trustworthy and human-centric AI within EU institutions. This role is further reinforced through international cooperation and cross-border exchanges of regulatory experience aimed at strengthening supervisory capacities and promoting convergence of best practices in AI oversight.³⁷⁷ Nevertheless, despite the expansion of its mandate, the EDPS remains primarily a supervisory and regulatory authority, as it neither adjudicates viola-

the free movement of such data, and repealing Directive 95/46/EC, 4 May 2016, OJ L 119/1, Arts. 68–76.

³⁷⁵ *Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L 2023/2854, Art. 37.*

³⁷⁶ *AI Act, Art. 57, 65, 70, 74 (9), 100.*

³⁷⁷ *European Data Protection Supervisor: Towards trustworthy AI in the EU public administration: The EDPS Compass for its new role under the AI Act (2026–2027)*, 2026, pp. 9–16. Online: https://www.edps.europa.eu/system/files/2026-03/26-03-17_edps-compass-under-ai-act-2026-2027_en.pdf.

tions of fundamental rights nor provides individual remedies comparable to those available before courts or specialised human rights mechanisms.

A recent study prepared for the European Parliament's Committee on Industry, Research and Energy examining the relationship between the AI Act and the broader EU digital legislative framework points to a number of structural and institutional challenges within the current system of AI governance. In particular, the study highlights the increasing complexity and fragmentation of the existing regulatory architecture, which may hinder coherent supervision, effective coordination between competent authorities, and the consistent enforcement of fundamental rights standards across the EU. These findings further illustrate some of the broader institutional limitations and governance gaps characterising the current framework for the oversight of AI systems.

Among its immediate priorities, the study underlines the importance of strengthening cooperation and coordination between supervisory authorities, particularly through the adoption of joint interpretative guidance, greater alignment of enforcement practices, and broader recognition of existing assessment procedures, including data protection impact assessments and fundamental rights impact assessments. In the medium term, the study advocates more targeted legislative reforms intended to clarify the distribution of institutional competences, minimise overlapping regulatory requirements, and improve the effective protection and enforceability of rights in AI-related contexts. From a broader structural perspective, the study ultimately points towards the need for a more integrated and coherent model of digital governance within the EU.³⁷⁸ In this regard, it emphasises simplification, institutional coherence, and strategic coordination as essential elements of a future regulatory framework capable of accommodating technological innovation while maintaining a high level of protection of fundamental rights and safety standards.

Alongside the supervisory and regulatory functions performed by the EDPB and the EDPS, the protection of fundamental rights in the context of AI also relies on institutions capable of identifying emerging risks and

³⁷⁸ GRAUX, H., GARSTKA, K., MURALI, N., CAVE, J., BOTTERMAN, M.: *Interplay between the AI Act and the EU digital legislative framework*. European Parliament – Policy Department for Transformation, Innovation and Health Directorate-General for Economy, Transformation and Industry, PE 778.575 - October 2025, pp. 78–86. Online: [https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU\(2025\)778575_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU(2025)778575_EN.pdf).

providing evidence-based expertise. Within the EU framework, this role is primarily fulfilled by the European Union Agency for Fundamental Rights (hereinafter “FRA”), whose mandate is centred on assisting EU institutions and Member States in ensuring compliance with fundamental rights standards in the implementation of EU law.³⁷⁹ Rather than exercising direct supervisory or enforcement powers, the FRA contributes through research activities, data collection, comparative analyses, and the development of expert opinions concerning the impact of digital technologies on fundamental rights. In the context of AI governance, its importance has further increased through its designation as a permanent member of the advisory forum established under Art. 67 of the *AI Act*, where it contributes specialised expertise relevant to the implementation and development of the EU regulatory framework on AI. Although the FRA does not function as an enforcement authority, its activities strengthen the broader institutional framework for the protection of fundamental rights by providing an important evidentiary and analytical basis for legislative, regulatory, and judicial responses to AI-related risks.³⁸⁰

The regulatory and institutional framework governing AI within the EU is characterised by a highly interconnected system of actors operating at both the national and supranational levels. Rather than relying on a single centralised authority, the EU approach is built upon a combination of Union-wide institutions, specialised agencies, national supervisory bodies, courts, and administrative authorities that collectively contribute to the oversight and enforcement of rules applicable to AI systems. This governance model reflects the broader constitutional structure of the EU, in which the protection of fundamental rights is implemented through continuous interaction between EU law and domestic legal systems.

Such a structure offers important advantages, particularly in enabling the adaptation of regulatory requirements to diverse national contexts and legal traditions. At the same time, however, the multiplicity of actors involved in

³⁷⁹ Council Regulation (EC) No. 168/2007 of 15 February 2007 establishing a European Union Agency for Fundamental Rights, OJ L 4M, 8.1.2008, Art. 2.

³⁸⁰ See e.g. Fundamental Rights Agency: *Assessing High-risk Artificial Intelligence: Fundamental Rights Risks*, 4 December 2025. Online: <https://fra.europa.eu/en/publication/2025/assessing-high-risk-ai>; Fundamental Rights Agency: *Bias in algorithms – Artificial intelligence and discrimination*, 8 December 2022. Online: <https://fra.europa.eu/en/publication/2022/bias-algorithm>; Fundamental Rights Agency: *Getting the future right – Artificial intelligence and fundamental rights*, 14 December 2020. Online: <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>.

AI governance may create practical difficulties concerning the coordination of competences, consistency of enforcement, and uniform interpretation of applicable standards. Differences in institutional capacities and regulatory approaches among Member States may further contribute to uneven levels of protection and legal uncertainty. In this respect, the EU framework illustrates the broader challenge of ensuring coherent and effective protection of fundamental rights within an increasingly complex and technologically driven regulatory environment.

9.2. Human Rights Protection and Emerging AI Governance within the Council of Europe

The COE has developed a particularly sophisticated regional framework addressing the relationship between AI and human rights. Building on its well-established normative foundations grounded in the ECHR, the organisation has progressively adapted existing human rights standards to respond to the challenges posed by technological developments.³⁸¹ To date, the Court has not directly addressed the issue of human rights violations arising specifically from AI systems.³⁸² Nevertheless, in cases concerning content moderation³⁸³ and facial recognition technologies,³⁸⁴ it has emphasised the obligation of States to comply with the requirements of proportionality, procedural fairness, and the prevention of arbitrary decision-making.³⁸⁵ Particular attention may be drawn to the judgment in *Glukhin v. Russia*, in which the Court examined the use of facial recognition technology by Russian authorities to identify and detain peaceful demonstrators. The Court concluded that the use of this technology in the circumstances of the case constituted a violation of both the right to respect for private and family life under Art. 8 ECHR

³⁸¹ BÄRDSEN, A.: *AI and Human Rights at the European Court of Human Rights*, Verfassungsblog, 8 December 2025. Online: <https://verfassungsblog.de/ai-and-human-rights-at-the-european-court-of-human-rights/>.

³⁸² SZAPPANYOS, M.: *Artificial Intelligence: Is the European Court of Human Rights Prepared?* In: *Acta Humana*, Vol. 93, No. 1, 2023, p. 98.

³⁸³ *Sanchez v. France*, Application no. 45581/15, Judgment of 15 May 2023; *Delfi AS v. Estonia*, Application no. 64569/09, Judgment of 16 June 2015; *Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary*, Application no. 22947/13, Judgment of 2 February 2016.

³⁸⁴ *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023.

³⁸⁵ PAPANASTASIOU, T. N., PAVLIDIS, G.: *The European Court of Human Rights and the Age of Artificial Intelligence: Jurisprudential Foundations for Future Adjudications*. In: *Cyprus Law Review*, Vol. 2–3, 2024, p. 187.

and the right to freedom of expression under Art. 10 ECHR. In its reasoning, the Court expressed serious doubts as to whether the domestic legal provisions authorising the processing of biometric personal data, including data obtained through facial recognition technology “in connection with the administration of justice”, satisfied the requirement of the “quality of law”, since they were formulated in excessively broad terms and appeared to permit the processing of such data in relation to virtually any type of judicial proceedings. The Court stressed that, for the “quality of law” requirement to be fulfilled in the context of facial recognition technologies, there must be detailed rules governing the scope and manner of application of such measures, accompanied by robust safeguards against the risks of abuse and arbitrariness. The need for such safeguards is even more pressing in cases involving live facial recognition technologies.³⁸⁶ The judgment therefore suggests that, in assessing future interferences arising from the use of AI systems, the Court is likely to place particular emphasis on the existence of a clear legal framework, adequate procedural safeguards, and a strict proportionality assessment capable of preventing arbitrary uses of such technologies.

The Court has further recognised the particularly sensitive nature of biometric data and facial recognition technologies in light of their potential impact on private life and individual autonomy. In *Gaughran v. the United Kingdom*, the Court confirmed that the retention of custody photographs, which will be held indefinitely on a local database for use by the police and that the police may also apply facial recognition and facial mapping techniques to the photograph constituted an interference with Article 8 ECHR.³⁸⁷ Likewise, in *S. and Marper v. the United Kingdom*, the Court warned that the use of advanced technological techniques in criminal justice systems cannot occur “at any cost” and must remain subject to careful balancing between public interests and the protection of private life.³⁸⁸

In its case law, the ECtHR has repeatedly acknowledged that contemporary security threats and technological developments may justify the use of increasingly sophisticated surveillance and data-processing tools by public authorities. At the same time, however, the Court has emphasised that the expansion of technological capabilities must be accompanied by a cor-

³⁸⁶ *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023, para. 82–83.

³⁸⁷ *Gaughran v. the United Kingdom*, Application no. 45245/15, Judgment of 13 February 2020, para. 70.

³⁸⁸ *S. and Marper v. the United Kingdom*, Application nos. 30562/04 and 30566/04, Judgment of 4 December 2008, para. 112.

responding strengthening of legal safeguards capable of protecting individuals against arbitrariness and disproportionate interferences with Convention rights. This approach is particularly visible in *Breyer v. Germany* and *Szabó and Vissy v. Hungary*, where the Court recognised both the legitimacy of adapting investigative techniques to modern communication technologies and the necessity of maintaining effective safeguards, oversight mechanisms, and proportionality requirements.³⁸⁹ Similar concerns were reflected in *Podchasov v. Russia*, in which the Court highlighted the risks associated with measures undermining encryption and secure digital communications, stressing that weakened encryption mechanisms may create broader systemic vulnerabilities affecting the privacy and security of all users.³⁹⁰

A recurring theme in the Court's jurisprudence concerns the requirement that technologically enabled interferences with privacy must be governed by sufficiently clear, foreseeable, and accessible legal frameworks containing robust safeguards against abuse and arbitrariness. In *Roman Zakharov v. Russia*, the Court criticised the absence of adequate oversight, insufficient limitations on surveillance powers, and the broad discretion granted to authorities in the context of secret surveillance measures.³⁹¹

More broadly, the Court's jurisprudence reflects growing concern regarding the implications of large-scale technologically enabled surveillance for democratic governance and the rule of law. In *Szabó and Vissy v. Hungary*, the Court warned against the dangers of "unfettered executive power" arising from uncontrolled surveillance techniques capable of intruding deeply into individuals' private spheres. These considerations are particularly relevant in the context of AI-driven surveillance and automated data-processing systems, whose scale, opacity, and predictive capacities may significantly intensify existing risks relating to privacy, autonomy, and arbitrary state interference.

A further important development within the Convention system is the advisory opinion mechanism introduced by Protocol No. 16 to the ECHR, which enables the highest national courts to request advisory opinions from the ECtHR on questions concerning the interpretation and application of Convention rights. Although not specifically designed for disputes involving

³⁸⁹ *Breyer v. Germany*, Application no. 50001/12, Judgment of 30 January 2020, para. 88; *Szabó and Vissy v. Hungary*, Application no. 37138/14, Judgment of 6 June 2016, para. 68.

³⁹⁰ *Podchasov v. Russia*, Application no. 33696/19, Judgment of 13 February 2024, para. 77–79.

³⁹¹ *Roman Zakharov v. Russia*, Application no. 47143/06, Judgment of 4 December 2015, para. 302–305.

AI systems, this mechanism may become increasingly relevant in cases raising novel legal questions related to algorithmic decision-making, biometric technologies, or AI-enabled surveillance. By strengthening judicial dialogue between domestic courts and the ECtHR, advisory opinions may contribute to the gradual clarification and harmonisation of Convention standards applicable to emerging technological contexts.

Against this background, it is important to consider not only treaty-based and institutionalised mechanisms, but also the role of other actors within the COE system that contribute to the protection and promotion of human rights in the context of technological change. In this regard, the COE Commissioner for Human Rights occupies a distinct position, operating through monitoring, dialogue, and soft influence rather than formal supervisory powers. Acting within this mandate, the Commissioner supports the effective protection of human rights primarily through the issuance of recommendations, thematic analyses, and continuous engagement with national authorities. At the same time, the Commissioner plays an important role in detecting systemic deficiencies in both legal frameworks and their practical implementation, while encouraging corrective measures. This role is complemented by cooperation with national human rights institutions, interaction with international actors, and regular reporting to the Committee of Ministers and the Parliamentary Assembly of the COE. Through these activities, the Commissioner contributes to strengthening the overall human rights framework by means of non-binding, yet influential, forms of intervention.³⁹²

Collectively, the outlined institutional framework combines judicial interpretation and supervision with advisory, monitoring and coordination-oriented mechanisms. The ECtHR nevertheless remains the principal judicial authority for assessing potential human rights violations arising from the development and use of AI systems.

9.2.1. Governance and Oversight under the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law

The practical effectiveness of the Framework Convention is closely linked not only to the substantive standards it establishes, but also to the institutional and procedural mechanisms designed to secure their implementation and supervision in practice. The implementation and oversight of the Framework Convention are entrusted to the Conference of the Parties, composed

³⁹² Council of Europe – Committee of Ministers: *Resolution (99) 50 on the Council of Europe Commissioner for Human Rights*, 7 May 1999, Art. 1–5.

of representatives of the Parties and mandated to facilitate implementation, address emerging issues, consider amendments, promote cooperation, and exchange information on legal and technological developments.³⁹³ This design choice reflects both the continued reliance on existing human-rights mechanisms and the underlying uncertainty as to their capacity to effectively address AI-related harms. At the same time, it suggests that further institutional adaptation may become necessary as the Convention is implemented in practice.³⁹⁴ The flexibility of the Convention's follow-up mechanism lies in the absence of a fixed timetable for its convocation. The Conference of the Parties is convened by the Secretary General of the COE as appropriate and periodically, or upon request of a majority of the Parties or the Committee of Ministers.³⁹⁵ In accordance with Art. 26, Parties must submit periodic reports on their implementation activities and engage in international cooperation to share information, assist non-Parties, and mitigate risks to human rights, democracy, and the rule of law.

To ensure compliance, Parties are required to establish or designate independent and impartial oversight mechanisms with adequate powers, expertise, and resources. These mechanisms must cooperate effectively, including with existing human rights structures where applicable.³⁹⁶ Given the pervasive use of AI systems across multiple sectors, as well as the fact that all States Parties already operate various regulatory and supervisory mechanisms for the protection of human rights, the provision emphasises the need to review existing mechanisms so that they can effectively apply within the context of activities taking place throughout the lifecycle of AI systems. States Parties may therefore choose to expand, transfer, modify, or redefine the functions of such mechanisms or, where appropriate, establish entirely new structures or bodies. Art. 26 expressly leaves these decisions to the discretion of the States Parties. While this approach offers greater flexibility, it may simultaneously result in divergent implementation practices among the Parties. Nevertheless, the practical functioning of these supervisory mechanisms has not yet been tested in practice, given the current absence of sufficient number of ratifications.

³⁹³ *Framework Convention*, Art. 23

³⁹⁴ *Ibid*, Art. 23 (2) (a) and (b).

³⁹⁵ *Explanatory Report to the Framework Convention*, point 131.

³⁹⁶ *Framework Convention*, Art. 26.

9.3. The African Institutional Framework for Human Rights and Artificial Intelligence

In light of the current stage of development of AI governance within the AU, no specialised regional regulatory or supervisory body has yet been established specifically for AI. This institutional absence, however, does not mean that the protection of human rights in the context of AI falls outside the existing African human rights system. On the contrary, AI-related harms may, at least in principle, be addressed through the mandates of existing regional human rights bodies, particularly the African Commission on Human and Peoples' Rights and the African Court on Human and Peoples' Rights. The central question is therefore not whether an AI-specific institutional mechanism currently exists, but whether existing human rights bodies are capable of interpreting and applying established guarantees to technologically mediated forms of interference with human rights.

The African Commission may play an important role in this regard through its promotional and protective mandate, including the examination of state reports, individual communications, resolutions, studies and thematic work. Its potential contribution lies primarily in the clarification of how rights protected under the African Charter apply to AI-related risks, including discriminatory decision-making, digital exclusion, surveillance, restrictions on freedom of expression, and the use of data-driven systems in access to public services. This potential is already reflected in *Resolution 473*, in which the Commission called for a study on human and peoples' rights and AI, robotics and other new and emerging technologies in Africa. The significance of this initiative lies not in the creation of a new regulatory regime, but in the recognition that AI governance should be assessed within the existing African human rights framework.

Although the African Commission has not yet decided cases concerning AI, its existing practice demonstrates an interpretative potential that may be relevant for future technology-related claims. In *Social and Economic Rights Action Center and Center for Economic and Social Rights v. Nigeria*, the African Commission clarified that the rights protected under the African Charter generate several layers of State obligations, including the duties to respect, protect, promote and fulfil human rights. These duties apply to both civil and political rights as well as economic, social and collective rights, and entail a combination of negative and positive obligations. Particularly relevant for AI-related harms is the duty to protect, which requires States to protect

right-holders against interference by third parties through legislation, regulation and effective remedies. The Commission further emphasised that this duty includes the creation and maintenance of an effective legal and regulatory framework enabling individuals to freely realise their rights and freedoms.³⁹⁷

The African Court on Human and Peoples' Rights may also contribute to the development of AI-related human rights standards, particularly in cases involving surveillance, discrimination, restrictions on expression, access to remedies, or violations of dignity and privacy-related interests. Unlike the Commission, the Court delivers legally binding judgments. Nevertheless, its practical role remains limited by jurisdictional and access-related constraints, especially the requirement that States accept direct access for individuals and non-governmental organisations. These limitations are particularly significant in the context of AI, where effective protection requires timely access to remedies, technical expertise, and the ability to address complex forms of harm involving both public authorities and private actors.

The African institutional framework therefore illustrates both the potential and the limits of relying on existing human rights mechanisms in the age of AI. Its strength lies in the capacity of the African Commission and the African Court to interpret established human rights guarantees in light of new technological risks. Its limitations lie in the absence of a specialised AI-related human rights mechanism, the non-binding character of the Commission's decisions, restricted access to the Court, limited enforcement capacity, and fragmented coordination between human rights institutions, data protection authorities and emerging digital governance bodies. For these reasons, the practical potential of the African Charter remains only partially realised in the context of AI.³⁹⁸ Strengthening the African regional response to AI will therefore depend not only on the development of AI-specific policies, but also on the integration of human rights institutions into broader digital governance frameworks.

³⁹⁷ *Social and Economic Rights Action Center (SERAC) and Center for Economic and Social Rights (CESR) / Nigeria* – 155/96, African Commission on Human and Peoples' Rights, 27 October 2001, para. 44–46 and 57–58.

³⁹⁸ FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: *African Journal of Legal Research*, Vol. 4, No. 2, 2026, p. 77.

9.4. The Inter-American Institutional Framework for Human Rights and Artificial Intelligence

As outlined in previous chapters, the Inter-American human rights system has not established a specialised body dealing exclusively with the human rights implications of AI. This institutional absence, however, does not render the system irrelevant for AI governance. On the contrary, the Inter-American Commission on Human Rights and the Inter-American Court of Human Rights may contribute to the protection of human rights in the context of AI through their existing mandates, particularly by interpreting established rights in light of new forms of surveillance, data processing, profiling and automated decision-making.

The Inter-American Court of Human Rights is especially relevant in this regard due to its evolutive interpretation of the American Convention on Human Rights. A particularly important development is the judgment in *Members of the José Alvear Restrepo Lawyers' Collective v. Colombia*, in which the Court recognised the right to informational self-determination as an autonomous human right. The case concerned, inter alia, surveillance, harassment and the collection of personal information by State intelligence agencies.³⁹⁹ Although it did not concern AI systems directly, the judgment is highly relevant for AI governance because it addresses the legal consequences of the collection, retention, control and use of personal data by public authorities. In this judgment, the Court held that the right to informational self-determination is grounded in the protective content of the American Convention, particularly in the rights to privacy and freedom of expression, as well as in the right to judicial protection. The Court further emphasised that this right functions not only as an autonomous right, but also as a guarantee for the protection of other rights, including privacy, honour, reputation and human dignity. Importantly, the Court indicated that the right extends to personal data held by public bodies and may also be relevant in relation to records or databases managed by private entities.⁴⁰⁰

The significance of this reasoning for AI lies in the fact that AI systems frequently depend on large-scale collection, processing and analysis of personal data. In many cases, such systems generate inferences about individuals, classify them into categories, or influence decisions affecting access to rights,

³⁹⁹ *Members of the José Alvear Restrepo Lawyers' Collective v. Colombia*, Inter-American Court of Human Rights, Judgment of 18 October 2023, Series C No. 506.

⁴⁰⁰ *Ibid.*, para. 588.

services and opportunities. The recognition of informational self-determination therefore provides an important human rights basis for requiring individuals to have meaningful control over data concerning them, particularly where data-driven systems are used by public authorities or where private actors exercise significant influence over the enjoyment of human rights.

The Court also connected the effectiveness of informational self-determination with the existence of adequate institutional and procedural safeguards. It held that States must provide adequate, swift, free and effective mechanisms or procedures enabling individuals to access and control personal data held about them.⁴⁰¹ This requirement is closely linked to the obligation to adopt domestic legal provisions under Art. 2 of the American Convention and entails not only the formal recognition of rights, but also the creation of practical mechanisms through which those rights can be exercised.⁴⁰² In the context of AI, this reasoning may support procedural guarantees relating to access to data, correction, deletion, transparency, oversight and effective remedies against harmful data-driven practices.

In the operative part of the judgment, the Court found Colombia internationally responsible for the violation of the right to informational self-determination, recognised under Articles 11(2) and 13(1) of the American Convention,⁴⁰³ in conjunction with the obligations to respect and guarantee rights and to adopt domestic legal provisions under Articles 1(1) and 2. The Court also ordered measures aimed at enabling the victims to access information collected about them and to exercise control over such data, including correction, cancellation or deletion where appropriate.⁴⁰⁴ This aspect of the judgment is particularly important because it confirms that the protec-

⁴⁰¹ *Ibid*, para. 599.

⁴⁰² “Where the exercise of any of the rights or freedoms referred to in Article 1 is not already ensured by legislative or other provisions, the States Parties undertake to adopt, in accordance with their constitutional processes and the provisions of this Convention, such legislative or other measures as may be necessary to give effect to those rights or freedoms.”

⁴⁰³ Art. 11 (2): “No one may be the object of arbitrary or abusive interference with his private life, his family, his home, or his correspondence, or of unlawful attacks on his honor or reputation.”

Art. 13 (1): “Everyone has the right to freedom of thought and expression. This right includes freedom to seek, receive, and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing, in print, in the form of art, or through any other medium of one’s choice.”

⁴⁰⁴ *Members of the José Alvear Restrepo Lawyers’ Collective v. Colombia*, Inter-American Court of Human Rights, Judgment of 18 October 2023, Series C No. 506, para. 1013.

tion of personal data requires effective and accessible mechanisms, rather than merely formal or theoretical guarantees.

The Inter-American institutional framework therefore demonstrates how existing human rights mechanisms may contribute to AI governance even in the absence of AI-specific bodies. Its contribution lies primarily in the interpretative and remedial potential of the Inter-American Commission and the Inter-American Court. The recognition of informational self-determination as an autonomous right may become particularly relevant for future AI-related claims involving surveillance, automated profiling, data retention, algorithmic decision-making or inadequate access to remedies. At the same time, the effectiveness of this framework will depend on domestic implementation, the existence of independent oversight bodies, and coordination between human rights institutions, data protection authorities and other actors involved in digital governance.

IV. EMERGING DIGITAL AND AI-RELATED RIGHTS: BEYOND THE EXPANSION OF EXISTING PROTECTION

The preceding analysis has demonstrated that existing human rights mechanisms and interpretative approaches remain capable of responding to many challenges arising from digitalisation and the use of AI, but also that certain technologically mediated harms expose limits in the scope, precision or practical effectiveness of existing protection. At the same time, the growing integration of digital and AI systems into social, economic and governmental structures has intensified debates on whether established human rights guarantees are sufficient, or whether new normative claims and safeguards are gradually emerging in technologically mediated environments.

This part of the monograph therefore focuses on the evolving discourse on digital and AI-related rights.⁴⁰⁵ It first examines digital rights as a broader category of emerging normative claims connected with the transformation of social life through digital technologies, data-driven infrastructures and networked communication. These claims do not form a settled or closed catalogue. Rather, they reflect attempts to preserve human dignity, autonomy, equality, participation, security and trust in the digital age.

Against this broader background, the analysis then turns to AI-related rights and safeguards as a more specific category of claims responding to the distinctive features of AI systems, including automation, opacity, prediction, algorithmic manipulation and the delegation of decision-making authority to computational systems. Particular attention is devoted to the question of whether such claims merely specify or extend existing human rights, or

⁴⁰⁵ See e.g. SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025. Online: <https://afp.oxford-aiethics.ox.ac.uk/sitefiles/white-paper-professor-yuval-shany.pdf>; ELAM, M., REICH, R.: *Stanford HAI Artificial Intelligence Bill of Rights: A White Paper for Stanford’s Institute for Human-Centered Artificial Intelligence*, January 2022. Online: <https://hai.stanford.edu/assets/files/stanford-hai-ai-bill-of-rights.pdf>; The White House Office of Science and Technology Policy: *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, October 2022. Online: <https://marketingstorageragrs.blob.core.windows.net/webfiles/Blueprint-for-an-AI-Bill-of-Rights.pdf>.

whether some of them point towards genuinely new forms of protection required by the realities of algorithmic governance.

10. EMERGING DIMENSIONS OF HUMAN RIGHTS IN THE DIGITAL AGE

The gradual process of digitalisation has given rise not only to reflections on the substantive evolution of existing human rights, but also to debates on the possible emergence of entirely new rights. Although AI has only recently become a central focus of legal and policy debates, many of the challenges it poses to human rights are rooted in earlier stages of digitalisation.⁴⁰⁶ These challenges were already reflected in the development of so-called digital rights, which addressed the implications of data-driven and networked technologies for fundamental rights. In this respect, the human rights implications of AI can be understood as a continuation and intensification of these earlier developments.

Before turning to so-called emerging AI-related human rights, it is necessary to first clarify what is meant by digital rights.⁴⁰⁷ As is the case with AI itself, there is no universally accepted definition of digital rights. This is partly understandable, given the rapid pace of digitalisation, which makes it difficult to capture all its dimensions and their implications for the emergence and exercise of digital rights. Building on this, parts of the scholarship suggest that digital rights may be understood as an emerging category within IHRL, reflecting both the reinterpretation of existing rights and the gradual development of new rights adapted to the digital environment.⁴⁰⁸ Such an ap-

⁴⁰⁶ IENCA, M., ANDORNO, R.: *Towards new human rights in the age of neuroscience and neurotechnology*. In: *Society and Policy*, Vol. 13, No. 5, 2017, pp. 1–27; Van EST, R., GERRITSEN, J.: *Human rights in the robot age Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality – Expert report written for the Committee on Culture, Science, Education and Media of the Parliamentary Assembly of the Council of Europe (PACE)*. The Hague: Rathenau Instituut, 2017. Online: <https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf>.

⁴⁰⁷ PETRYSHYN, O. V., HYLIKA, O. S.: *Human Rights in the Digital Age: Challenges, Threats and Prospects*. In: *Journal of the National Academy of Legal Sciences of Ukraine*, Vol. 28, No. 1, 2021, p. 16.

⁴⁰⁸ SHANY, Y.: *Digital Rights and the Outer Limits of International Human Rights Law*. In: *German Law Journal*, Special Issue: International Law and Digitalization, Vol. 24, 2023, pp. 461–472; BIELIAKOV, K. I., TYKHOMYROV, O. O., RADOVETSKA, L. V., & KOSTENKO, O. V.: *Digital rights in the human rights system*. In: *InterEULawEast: Journal for the International and European Law, Economics and Market Integrations*, Vol. 10, No. 1, 2023, pp. 183–207.

proach seeks to preserve the philosophical core of human rights—grounded in freedom, dignity and autonomy—while adapting their content and modes of protection to the specificities of the digital environment. This need for reconceptualisation stems from the fact that traditional understandings of key rights were developed in a predominantly physical context and do not always fully capture the complexities and risks arising in digital spaces.⁴⁰⁹

Against this background, it becomes necessary to turn to existing doctrinal approaches that seek to define digital rights and to outline their essential characteristics. Digital rights are often broadly understood as encompassing the rights of individuals to access, use, create, and disseminate digital works, as well as to access computers, other electronic devices, and communication networks, most notably the Internet.⁴¹⁰ However, such expansive definitions have been subject to criticism. *Alsharif and Hnit*, for instance, argue that any meaningful conceptualisation of digital rights must also incorporate their negative dimension, including obligations to refrain from intrusive surveillance and from the commodification of personal data.⁴¹¹

This broader understanding is also reflected at the policy level, for example in the *European Declaration on Digital Rights and Principles for the Digital Decade*, which reinforces the idea that digital rights encompass both enabling and protective dimensions of human rights in the digital environment. This dual character is evident, in particular, in the recognition of rights related to access to high-quality connectivity and digital services,⁴¹² as well as in the emphasis on the protection of privacy and personal data, safeguards against unlawful surveillance, and the promotion of a safe and secure digital environment.⁴¹³ The Declaration further presents AI as a human-centred tool designed to enhance well-being, while emphasising the importance of

⁴⁰⁹ MORSKA, N.: *The Relationship Between Natural and Digital Human Rights: Problems of Legal Definition*. In: *Philosophy and Governance*, Vol. 18, No. 2, 2026, p. 3; CUSTERS, B.: *New digital rights: Imagining additional fundamental rights for the digital era*. In: *Computer Law & Security Review*, Vol. 44, 2022, 105636, p. 4.

⁴¹⁰ BIELIAKOV, K. I., TYKHOMYROV, O. O., RADOVETSKA, L. V., & KOSTENKO, O. V.: *Digital rights in the human rights system*. In: *InterEULawEast: Journal for the International and European Law, Economics and Market Integrations*, Vol. 10, No. 1, 2023, p. 186.

⁴¹¹ ALSHARIF, M., HNIT, H.: *Digital human rights: Legal debates and emerging foundations under the international bill of human rights*. In: *Social Sciences & Humanities Open*, Vol. 12, 2025, 102160, p. 1.

⁴¹² European Parliament, Council of the European Union, European Commission, Directorate-General for Communications Networks, Content and Technology: *European Declaration on Digital Rights and Principles for the Digital Decade*, OJ C 23, 23.1.2023, Art. 3. and 7.

⁴¹³ *Ibid*, Art. 16–22.

informed user choice, transparency, non-discrimination, human oversight, and safeguards ensuring the protection of fundamental rights.⁴¹⁴ This approach reflects a broader normative shift, according to which digital technologies are no longer perceived merely as objects of regulation, but as structural factors shaping the conditions for the exercise of human rights.

Therefore, in the context of pervasive technological development and the increasing digitalisation of social, economic and institutional life, digital rights must be regarded as an integral part of the contemporary international human rights framework rather than as optional or peripheral guarantees. Future legal instruments and interpretive practices should reflect this reality to ensure the full and inclusive protection of human dignity in an increasingly interconnected world.⁴¹⁵ Nevertheless, the recognition of digital rights as fundamental human rights remains a contested issue. Critics urge caution against premature legal codification due to the evolving nature of technology, while proponents emphasize the indispensability of digital access and protections in modern life.⁴¹⁶

Notwithstanding these conceptual disagreements, contemporary scholarly and expert debates increasingly point to the emergence of a broad spectrum of new digital rights and interests requiring legal recognition and protection in the context of digital technologies, including AI. From these discussions,⁴¹⁷ it is possible to identify, *inter alia*, the right to internet access and its counterpart, the right to be offline or disconnect, the right to digital identity, the right to cybersecurity, the right to digital privacy and the protection of personal data, the right to digital anonymity, the right to

⁴¹⁴ *Ibid*, Art. 8–9.

⁴¹⁵ ALSHARIF, M., HNIT, H.: *Digital human rights: Legal debates and emerging foundations under the international bill of human rights*. In: Social Sciences & Humanities Open, Vol. 12, 2025, 102160, p. 9.

⁴¹⁶ *Ibid*.

⁴¹⁷ MALGIERI, G., CUSTERS, B.: *Pricing Privacy – the right to know the value of your personal data*. In: Computer Law and Security Review, Vol. 34, No. 2, 2018, pp. 289–303. BIELIAKOV, K. I., TYKHOMYROV, O. O., RADOVETSKA, L. V., & KOSTENKO, O. V.: *Digital rights in the human rights system*. In: InterEULawEast: Journal for the International and European Law, Economics and Market Integrations, Vol. 10, No. 1, p. 193; ALSHARIF, M., HNIT, H.: *Digital human rights: Legal debates and emerging foundations under the international bill of human rights*. In: Social Sciences & Humanities Open, Vol. 12, 2025, 102160, p. 3–4; SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025; CUSTERS, B.: *New digital rights: Imagining additional fundamental rights for the digital era*. In: Computer Law & Security Review, Vol. 44, 2022, 105636, pp. 6–12.

be forgotten,⁴¹⁸ the right to a “clean” digital slate or digital reintegration, the right to know the value of one’s data, as well as rights related to freedom of digital content creation, digital heritage, and protection against misinformation and algorithmic manipulation.

The emergence of digital rights therefore reflects a broader transformation of human rights protection in technologically mediated environments. These rights do not form a closed or settled catalogue. Rather, they express a set of evolving normative claims aimed at preserving human dignity, autonomy, equality, participation and security in the digital age. Some of them may be understood as digital specifications of existing rights, while others point towards new interests that are not fully captured by traditional human rights categories.

In the context of emerging dimensions of human rights in the digital age, contemporary scholarly debates have also begun to explore whether AI systems or robots could eventually be recognised as holders of certain rights.⁴¹⁹ While such debates raise important theoretical questions concerning legal personality and the future development of law, they do not concern human rights in the strict sense. The present monograph focuses on the protection of human beings affected by AI and digital technologies. Questions concerning the possible rights of AI systems or robots therefore fall outside the scope of the analysis.

The relevant question for the present purposes is therefore not whether AI systems may themselves become rights-holders, but whether human beings require new or more specific rights and safeguards in response to the risks created by AI systems. On this basis, AI-related rights and safeguards may be understood as a more specific set of normative claims responding to the distinctive risks created by algorithmic systems, automated decision-making and increasingly data-driven forms of governance.

10.1. AI-Related Rights and Safeguards

Shany’s proposal for a future AI Bill of Human Rights provides a useful starting point for identifying emerging AI-related rights and safeguards. His White Paper sets out an initial list of seven rights aimed at ensuring that AI

⁴¹⁸ See Art. 17 of GDPR.

⁴¹⁹ DRACHOVSKÁ, K.: *Umělá inteligence jako nositelka základních práv?* In: *Právník*, Vol. 4, 2021, pp. 273–284; MÜLLER, V.C.: *Is it time for robot rights? Moral status in artificial entities.* In: *Ethics and Information Technology*, Vol. 23, 2021, pp. 579–587.

systems are developed and used in a manner compatible with human well-being, the minimisation of harm, and standards of fairness and justice. These include:

- (1) the right of access to AI systems,
- (2) privacy-related protections against harmful uses of AI,
- (3) the right to be free from algorithmic bias and unfairness,
- (4) the right to algorithmic transparency and explainability,
- (5) the right not to be subjected to algorithmic manipulation,
- (6) the right to a human decision and human-to-human interaction, and
- (7) the right to accountability for harms caused by the use of AI systems.⁴²⁰

Similar concerns are reflected in other AI bill of rights initiatives. The *Stanford HAI proposal* and the *White House Blueprint for an AI Bill of Rights* do not necessarily formulate rights in the same way as Shany, but they point to comparable protective concerns, namely, safety, non-discrimination, privacy, transparency, explanation, contestability, human alternatives and remedies.⁴²¹ These instruments therefore confirm that AI-related rights and safeguards are emerging around a relatively stable set of normative concerns, even if their legal status and terminology remain unsettled. What remains contested, however, is whether these concerns should be understood merely as policy objectives and ethical safeguards, or whether some of them are sufficiently fundamental to warrant recognition in human rights terms. Shany's argument for an AI Bill of Human Rights addresses this issue directly. It is based on the claim that certain aspects of AI systems affect basic human needs and interests in ways that may justify the development of new human rights protections.⁴²² This is particularly the case where such interests are closely connected to the moral foundations of human rights, including dignity, liberty and solidarity, and where they are not adequately covered by existing human rights guarantees. Existing human rights law may offer partial protection in certain contexts, for example through due process guarantees or the prohibition of inhuman or degrading treatment, but it does not neces-

⁴²⁰ SHANY, Y.: *White Paper - The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, p. 70.

⁴²¹ ELAM, M., REICH, R.: *Stanford HAI Artificial Intelligence Bill of Rights: A White Paper for Stanford's Institute for Human-Centered Artificial Intelligence*, January 2022; The White House Office of Science and Technology Policy: *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, October 2022.

⁴²² SHANY, Y.: *White Paper - The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, pp. 42–45.

sarily provide comprehensive protection for the full range of dignity, fairness and agency-related concerns raised by AI systems.⁴²³

At the same time, the proposed framework should not be understood as implying that all these claims necessarily amount to genuinely new human rights. A significant part of emerging rights may be interpreted as a specification or extension of already established rights and principles within IHRL.⁴²⁴ For example, privacy-related protections against harmful uses of AI systems are closely connected to the existing right to privacy and the protection of personal data. Similarly, the right to be free from algorithmic bias and unfairness largely builds upon the principles of equality and non-discrimination, while the right to accountability for harms caused by AI systems may be linked to the right to an effective remedy and broader principles of responsibility and due diligence. Even algorithmic transparency and explainability can, in many contexts, be understood as procedural safeguards supporting the effective exercise of existing rights, particularly in relation to due process, access to information and the ability to challenge decisions affecting one's legal position or interests.

The recognition of a genuinely new human right should nevertheless remain exceptional. Technological novelty alone cannot justify the proliferation of new rights. A distinct AI-related right appears warranted only where the interest requiring protection is of fundamental importance to human dignity, autonomy or equality, where existing human rights do not provide sufficiently clear or effective protection, and where the proposed right can be formulated with sufficient normative precision to generate identifiable obligations and safeguards. Where these conditions are not met, adaptation through interpretation, procedural guarantees or more specific regulatory duties may offer a more coherent response within the existing human rights framework.

Considering this distinction between the extension of existing rights and the emergence of more specific AI-related rights, the following subchapters focus on two selected examples: the right not to be subjected to algorithmic manipulation and the right to a human decision and human-to-human interaction. These two claims are selected because they most clearly reveal the limits of analysing AI-related harms solely through established rights such as privacy, equality, due process or effective remedies. Algorithmic manipula-

⁴²³ *Ibid*, p. 39.

⁴²⁴ CUSTERS, B.: *New digital rights: Imagining additional fundamental rights for the digital era*. In: Computer Law & Security Review, Vol. 44, 2022, 105636, pp. 6–12.

tion may undermine mental integrity, conscious deliberation and individual agency, while the replacement of human decision-making and interaction by AI systems raises dignity-related concerns about treating individuals as full human beings rather than as data-driven profiles.⁴²⁵ In this sense, both claims go beyond the correction of defective AI outputs and address the more fundamental question of how human autonomy, dignity and interpersonal relations can be preserved in AI-mediated environments.

10.1.1. Right Not to Be Subjected to Algorithmic Manipulation

The proposed right not to be subjected to algorithmic manipulation addresses one of the most distinctive risks of AI-mediated environments: the capacity of algorithmic systems to influence human behaviour, preferences and decision-making processes in ways that may bypass conscious deliberation and undermine individual autonomy.⁴²⁶ AI systems create conditions in which it becomes increasingly difficult to distinguish authentic from inauthentic content, including misinformation, disinformation and deepfakes. Manipulation therefore does not arise only from explicit deception or false representations of reality. It may also result from subtle nudges, subliminal influences and personalised forms of behavioural steering that push individuals to act, or refrain from acting, in ways that bypass rational reflection and conscious deliberation. It directly affects individual agency and autonomy, and may interfere with freedom of thought, freedom of opinion, the right to seek and receive information and, in political contexts, the right to take part in public affairs.⁴²⁷ The concern with algorithmic manipulation becomes particularly serious where AI-mediated environments shape not only individual choices, but also the formation of personality, critical thinking and exposure to diverse information. In this respect, algorithmic manipulation may affect freedom of thought, expression and action by creating informational environments that reinforce existing preferences, reduce exposure to alternative viewpoints and make individuals more vulnerable to behavioural steering.

⁴²⁵ SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, pp. 13–14.

⁴²⁶ VAN GILS, F., MÜLLER, W., PRÜFER, J.: *Microtargeting, voters' unawareness, and democracy*. In: *The Journal of Law, Economics, and Organization*, Vol. 41, No. 2, 2025, p. 651; PARK, P. S., GOLDSTEIN, S., O'GARA, A., CHEN, M., HENDRYCKS, D.: *AI deception: A survey of examples, risks, and potential solutions*. In: *Patterns*, Vol. 5, No. 5, 2024, p. 2.

⁴²⁷ SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, p. 61.

This concern is especially acute in relation to children and young people, whose cognitive and emotional development may be affected by prolonged exposure to persuasive and attention-driven digital systems.⁴²⁸

These concerns are reflected in numerous soft law and regulatory instruments on AI. Several of them identify manipulation as a specific risk associated with AI systems, particularly where such systems exploit vulnerabilities, constrain free choices, disseminate misinformation or disinformation, or rely on manipulative interface design, including so-called dark patterns. In this respect, the concern with algorithmic manipulation overlaps with broader debates on informational integrity in the digital age.⁴²⁹ This concern is visible, for example, in the Global Digital Compact, which calls for strengthened international cooperation to address misinformation, disinformation, hate speech and information manipulation in a manner consistent with international law.⁴³⁰ Similar concerns are reflected in the European Declaration on Digital Rights and Principles, which calls for digital and cyber environments that protect individuals from manipulation,⁴³¹ and in the African Union's Continental AI Strategy, which emphasises the protection of informational integrity.⁴³² These instruments show that manipulation is increasingly understood not merely as an individual consumer harm, but as a broader threat to autonomy, public discourse and democratic participation. The most explicit legal formulation can be found in the AI Act, which prohibits certain AI practices involving subliminal techniques or purposefully manipulative or deceptive techniques where they materially distort the behaviour of a person or group by impairing their ability to make an informed decision and are likely to cause significant harm.⁴³³ This prohibition is important because it recognises that the harm of manipulative AI lies not only in the pro-

⁴²⁸ BRITO, I.: *The right not to be monitored: in defence of one of the last bastions of the human being against the emerging omniscient surveillance society*. In: UNIO – EU Law Journal, Vol. 11, No. 1, June 2025, pp. 79–80.

⁴²⁹ SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, pp. 13–14.

⁴³⁰ Global Digital Compact UN General Assembly: *Pact for the Future – Annex – Global Digital Compact*, A/RES/79/1, 22 September 2024, para. 33.

⁴³¹ European Parliament, Council of the European Union, European Commission, Directorate-General for Communications Networks, Content and Technology: *European Declaration on Digital Rights and Principles for the Digital Decade*, OJ C 23, 23.1.2023, Art. 15(d), 16(b).

⁴³² AU: *Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity*, July 2024., pp. 49–51

⁴³³ AI Act, Art. 5 (1).

duction of false or misleading information, but also in the impairment of autonomous decision-making. The same logic is reflected in the White House *Blueprint for an AI Bill of Rights*, which rejects user experience design choices that intentionally obscure or manipulate user choice, including dark patterns.⁴³⁴ Although these instruments differ in legal nature and terminology, they point to a common normative concern: AI systems should not be used to covertly steer, deceive or exploit individuals in ways that undermine human dignity, autonomy and the capacity for informed choice.

The proposed right not to be subjected to algorithmic manipulation should not be understood as an absolute prohibition of all forms of influence. Its scope is necessarily relative, since the distinction between legitimate and illegitimate manipulation is often difficult to draw. Some forms of influence may be subtle but socially beneficial, such as nudges aimed at improving health or safety, while others may cause only trivial harm or occur in contexts in which individuals knowingly consent to being influenced. The assessment of legitimacy should therefore depend on the means used, the purpose pursued, the degree of transparency and the harm caused.⁴³⁵ In this sense, the right would likely have to develop gradually through specific use cases, distinguishing permissible forms of influence from manipulative practices that undermine autonomy, informed choice and mental integrity.

Against this background, the protection against algorithmic manipulation may be understood as a particularly strong candidate for more specific human rights protection where existing guarantees relating to privacy, freedom of thought or expression do not sufficiently capture interference with the process of autonomous preference formation and decision-making itself.

10.1.2. The Right to a Human Decision and a Human-To-Human Interaction

The proposed right to a human decision and a human-to-human interaction reflects a growing concern that individuals should not be left solely to automated or AI-mediated processes in situations that significantly affect their rights, interests, opportunities or well-being.⁴³⁶ Unlike more established

⁴³⁴ The White House Office of Science and Technology Policy: *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, October 2022, p. 34.

⁴³⁵ SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, pp. 61–62.

⁴³⁶ FERREIRA, J. J., MONTEIRO, S.: *The human-AI relationship in decision-making: AI explanation to support people on justifying their decisions*. In: Joint Proceedings of the ACM

guarantees, this claim has not yet acquired a fixed legal meaning. It appears in different formulations across scholarly, policy and regulatory debates, including the right to a human decision, the right to human review, the right to opt out of automated decision-making, the right to human alternatives and fallback, meaningful human control, and the right to meaningful human contact.⁴³⁷

For instance, the White House *Blueprint for an AI Bill of Rights* formulates the principle of “human alternatives, consideration, and fallback”. According to this principle, individuals should, where appropriate, be able to opt out of automated systems in favour of a human alternative and should have access to timely human consideration and remedy where an automated system fails, produces an error, or where its impact is contested. The Blueprint further stresses that human consideration and fallback should be accessible, equitable and effective, and should not impose an unreasonable burden on the public. In sensitive domains, such as criminal justice, employment, education and health, automated systems should incorporate meaningful human oversight and human consideration for adverse or high-risk decisions.⁴³⁸ Thus, while a fully-fledged right to a human decision has not yet been generally recognised as an autonomous human right, its normative components are already present in several legal and policy frameworks. In particular, the right appears in a fragmented form in data protection and digital rights instruments, through safeguards against solely automated decision-making, the right to obtain human intervention, the right to express one’s point of view, the right to contest automated decisions and requirements of human

IUI 2021 Workshops, April 13-17, 2021, pp. 2–3. Online: <https://ceur-ws.org/Vol-2903/IUI21WS-TESS-7.pdf>.

⁴³⁷ See e.g. Van EST, R., GERRITSEN, J.: *Human rights in the robot age Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality – Expert report written for the Committee on Culture, Science, Education and Media of the Parliamentary Assembly of the Council of Europe (PACE)*. The Hague: Rathenau Instituut, 2017. The White House Office of Science and Technology Policy: *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, October 2022, p. 7; CHENG, M.: *The right to human decision: analyzing policies, ethics, and implementation*. First International Conference on Addressing Socioethical Effects of Artificial Intelligence, 2024, p. 3. Online: <https://openreview.net/pdf?id=VrxTvBw0ly>.

⁴³⁸ The White House Office of Science and Technology Policy: *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, October 2022, p. 7.

oversight or supervision. These elements can be found, *inter alia*, in Art. 22 GDPR, and the African Union's Malabo Convention.⁴³⁹

The recognition of a right to a human decision is supported by part of the scholarship on the grounds that where algorithmic systems make decisions that significantly affect individuals' lives, such as decisions concerning employment, access to credit, social benefits, compensation or legal responsibility, the absence of any possibility of appeal may raise serious concerns from the perspective of due process and fairness. The ability to contest a decision and to have one's case reconsidered is a core feature of legal protection, and algorithmic decisions should not become practically incontestable merely because they are produced by automated systems.

Second, a right to a human decision may strengthen accountability and human agency in relation to AI systems. Since algorithmic systems may produce erroneous, unfair or context-insensitive outcomes, individuals should have access to a mechanism capable of identifying and correcting such failures. This is particularly important because even systems that perform well on average may fail in individual cases. Human review may therefore serve as a safeguard against the risk that individuals are reduced to statistical profiles, scores or data points without meaningful consideration of their particular circumstances.⁴⁴⁰ The argument also has a dignity-based dimension. Shany contends that denying individuals access to human decision-making may have a dehumanising effect, especially where consequential decisions are made solely through algorithmic categorisation. On this view, the right to a human decision is not only a procedural safeguard, but also a means of ensuring that individuals are treated as human beings endowed with dignity rather than as members of algorithmically defined groups. This rationale connects the proposed right to broader concerns already familiar to IHRL, including arbitrariness, accountability, equality, political participation and human dignity.⁴⁴¹

The proposed right to a human decision and human-to-human interaction should not be understood as absolute. Its scope must be assessed in light

⁴³⁹ *African Union Convention on Cyber Security and Personal Data Protection*, Malabo, 27 June 2014, Art. 14 (5).

⁴⁴⁰ CHENG, M.: *The right to human decision: analyzing policies, ethics, and implementation*. First International Conference on Addressing Socioethical Effects of Artificial Intelligence, 2024, pp. 3–4.

⁴⁴¹ HUQ, A.: *A Right to A Human Decision*. In: *Virginia Law Review*, Vol. 106, 2020, p. 687; SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, p. 63.

of legality, legitimacy, necessity and proportionality. In routine contexts, such as the calculation of tax rates, parking fines or standard consumer complaints, reliance on automated systems may be acceptable, particularly where individuals have access to effective appeal mechanisms. In more sensitive domains, however, such as criminal justice or life-and-death decisions in health care, stronger human involvement may be required. The normative point of departure should be that individuals should, as far as possible, be able to opt out of AI-mediated decision-making or challenge automated decisions before a human being. Consent alone should not always justify exclusive reliance on algorithms, especially where serious harm may result or where the societal importance of human judgement is particularly high.⁴⁴²

The right to a human decision is not without difficulties. Its recognition does not automatically guarantee better, fairer or more ethical outcomes, since human reviewers may themselves be biased, inconsistent or unable to meaningfully assess algorithmic outputs. Moreover, the practical meaning of “meaningful” human review remains unclear, particularly where review mechanisms risk becoming a mere formal confirmation of algorithmic decisions. The effectiveness of such a right also depends on accessibility, promptness and equality of use; otherwise, it may exacerbate existing information and resource inequalities by benefiting primarily those individuals who are better equipped to contest automated decisions.⁴⁴³

These objections do not necessarily undermine the relevance of the proposed right, but they indicate that its content must be carefully differentiated according to context. The appropriate safeguard may range from human review or appeal, through a right to opt out in favour of a human alternative, to stronger limits on fully automated decision-making in areas where dignity, vulnerability, liberty or access to essential services are at stake.

The analysis of these two proposed rights confirms that the emergence of AI does not necessarily require the wholesale creation of a new catalogue of human rights. Many AI-related concerns can be addressed through the interpretation and specification of existing guarantees, supplemented by procedural and regulatory safeguards. At the same time, certain forms of algorithmic interference may affect interests, particularly autonomous deci-

⁴⁴² SHANY, Y.: *White Paper – The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025, pp. 63–65.

⁴⁴³ CHENG, M.: *The right to human decision: analyzing policies, ethics, and implementation*. First International Conference on Addressing Socioethical Effects of Artificial Intelligence, 2024, pp. 3–4.

sion-making, mental integrity and meaningful human agency, in ways that are not fully captured by existing rights. In such cases, the recognition of more specific AI-related rights may become justified where it provides distinct normative protection that cannot be achieved effectively through interpretation alone. The development of such rights should therefore remain cautious and context-sensitive, preserving the coherence of the existing human rights framework while allowing it to respond to genuinely new forms of technological power.

V. CONCLUSIONS AND FUTURE PERSPECTIVES

This monograph has examined the impact of AI on the protection of human rights from the perspective of IHRL. Its central question was whether the existing system of IHRL provides sufficient normative and institutional tools to address the risks associated with AI systems, or whether these developments require new legal approaches, new rights, or new institutional mechanisms at the international level.

The analysis has shown that IHRL remains the principal normative framework for assessing the development and deployment of AI systems. Despite the novelty, technical complexity and increasing autonomy of AI technologies, the risks they generate do not arise in a legal vacuum. They engage rights and principles already recognised within universal and regional human rights instruments, including human dignity, privacy, data protection, equality and non-discrimination, freedom of expression, fair trial guarantees, access to justice, and economic and social rights. These rights continue to provide an essential legal vocabulary for evaluating AI-related harms and for defining the obligations of States and, indirectly, the responsibilities of private actors.

At the same time, the analysis has demonstrated that AI challenges the traditional assumptions upon which IHRL has historically operated. IHRL developed primarily within a state-centred framework, in which violations were typically attributed to public authorities and addressed through territorially defined obligations. AI systems, by contrast, are often developed, deployed and controlled through complex networks of public and private actors. Their operation may be transnational, opaque, data-driven and difficult to attribute to a single decision-maker. These features expose structural limitations in existing human rights protection, particularly in relation to accountability, transparency, access to remedies and institutional capacity.

One of the main findings of this monograph is therefore that the issue is not the absence of relevant human rights norms, but rather the difficulty of applying and enforcing them effectively in technologically mediated environments. Existing human rights guarantees remain relevant, but they require dynamic interpretation and operationalisation. Rights such as privacy,

non-discrimination and fair trial guarantees must be understood in light of datafication, algorithmic classification, automated decision-making, profiling and large-scale surveillance. Similarly, principles such as human dignity, autonomy and equality acquire renewed significance in contexts where decisions affecting individuals may be shaped by systems that are difficult to understand, contest or control.

This conclusion is supported by the jurisprudence analysed in the monograph. Although courts have not yet developed a comprehensive AI-specific human rights doctrine, existing case law already provides important principles for the governance of AI and data-driven systems. The case law of the CJEU in *Digital Rights Ireland*, *Tele2 Sverige*, *La Quadrature du Net*, *Schrems I* and *Schrems II* confirms the importance of clear legal frameworks, proportionality, independent oversight, effective remedies and safeguards against indiscriminate data processing and surveillance. These principles are highly relevant for AI systems, particularly where they rely on large-scale data collection, profiling or predictive analytics. The CJEU's judgment in *SCHUFA Holding* is especially important because it adopts a functional understanding of automated decision-making under Article 22 GDPR and demonstrates that formally human decisions may still fall within the logic of automated decision-making where algorithmic scoring effectively determines the outcome in practice. Pending cases such as *Like Company v. Google Ireland* and *Yettel Bulgaria* further illustrate that courts will increasingly be called upon to address questions concerning generative AI, text and data mining, algorithmic transparency, access to source code or "black box" information, and effective judicial protection in automated decision-making contexts.

The case law of the ECtHR similarly remains relevant for AI governance, even if most cases were not decided in an AI-specific context. Judgments concerning surveillance, data retention, biometric data, online expression and digital monitoring, including *Big Brother Watch*, *Zakharov*, *S. and Marper*, *Glukhin* and *Podchasov*, demonstrate the Court's insistence on legality, foreseeability, necessity, proportionality and effective safeguards against abuse. These requirements provide a useful framework for assessing AI-enabled surveillance, facial recognition, predictive policing and other intrusive forms of data-driven governance. In the Inter-American system, the recognition of informational self-determination in *Members of the José Alvear Restrepo Lawyers' Collective v. Colombia* is particularly significant. It confirms that the protection of personal data requires effective and accessible mechanisms enabling individuals to access, control, correct or delete infor-

mation concerning them. This reasoning may become important for future AI-related claims involving surveillance, automated profiling, data retention, algorithmic decision-making and inadequate access to remedies.

The monograph has also shown that soft law has played a particularly important role in the development of AI governance. At the universal level, instruments and initiatives adopted within the United Nations system, UNESCO, UNICEF and the OECD have contributed to the articulation of common principles, including human rights, transparency, accountability, fairness, non-discrimination, safety, human oversight and the protection of vulnerable groups, which helped to consolidate a common normative vocabulary and have functioned as a laboratory for the progressive development of legal standards. Their importance lies not only in their immediate regulatory effect, but also in their ability to shape later binding instruments and institutional practice.

However, soft law alone cannot ensure effective human rights protection in the context of AI. The analysis of regional binding approaches, particularly the EU AI Act and the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, demonstrates a gradual shift from ethical principles and voluntary standards towards more structured legal and institutional frameworks. The AI Act represents the most detailed regulatory model currently adopted, combining risk classification, prohibited practices, obligations for high-risk AI systems, transparency duties, requirements concerning human oversight, and fundamental rights impact assessments in certain contexts. The Framework Convention, in turn, provides the first international treaty framework expressly linking AI governance with human rights, democracy and the rule of law. Together, these instruments indicate that AI governance is entering a phase of legal institutionalisation.

Nevertheless, the emergence of binding frameworks does not eliminate the risk of fragmentation. On the contrary, the coexistence of universal soft law, regional binding instruments, national regulatory regimes, technical standards and sector-specific rules may create a complex and uneven regulatory landscape. While regional innovation may contribute to the development of stronger safeguards, it may also lead to divergences in the level of human rights protection. This is particularly relevant in light of the cross-border nature of the AI system life cycle, where design, training, deployment and impact may occur across different jurisdictions. Future regulation will therefore require greater coordination between universal and regional

frameworks, as well as between human rights bodies, data protection authorities, AI regulators and other supervisory institutions.

The institutional dimension of human rights protection is thus of central importance. Traditional human rights bodies, including UN treaty bodies, special procedures and regional human rights mechanisms, remain relevant for AI-related harms. Their interpretative practice can clarify how existing human rights obligations apply to algorithmic systems and data-driven decision-making. At the same time, these bodies often lack the technical expertise, resources and procedural tools necessary to respond effectively to complex AI-related risks. This does not make them obsolete, but it does require institutional adaptation, cooperation with specialised regulators and stronger integration of technical expertise into human rights assessment.

The regional analysis confirms that different institutional pathways are emerging. In Europe, the EU and the Council of Europe have developed the most advanced institutional and regulatory responses, combining legal obligations with supervisory and governance mechanisms. The EU AI Act establishes a complex governance architecture involving national competent authorities, market surveillance authorities, the European Artificial Intelligence Board and other supporting bodies. The Council of Europe Framework Convention, complemented by the HUDERIA methodology, provides a human rights-oriented model for risk and impact assessment. HUDERIA is particularly important because it translates abstract commitments to human rights, democracy and the rule of law into a structured assessment process based on context-based risk analysis, stakeholder engagement, risk and impact assessment, mitigation planning and iterative review. In this respect, it demonstrates that human rights protection in the AI context cannot be reduced to abstract principles, but must be embedded in concrete institutional and procedural practices throughout the entire AI system life cycle.

In the African and Inter-American systems, by contrast, AI-specific regulatory bodies have not yet been established at the regional level. Nevertheless, existing human rights institutions, particularly the African Commission on Human and Peoples' Rights, the African Court on Human and Peoples' Rights, the Inter-American Commission on Human Rights and the Inter-American Court of Human Rights, may still contribute to AI governance through the interpretation of established human rights guarantees. The recognition of informational self-determination by the Inter-American Court illustrates how existing human rights mechanisms may generate concepts that are highly relevant for the protection of individuals in data-driven en-

vironments. At the same time, the effectiveness of these systems will depend on domestic implementation, institutional capacity, independent oversight and cooperation with data protection and digital governance authorities.

A further finding concerns the role of private actors. AI systems are frequently developed and deployed by companies whose activities may significantly affect the enjoyment of human rights. Although international human rights obligations remain primarily incumbent upon States, the growing influence of private technological actors makes the business and human rights framework increasingly important. The UN Guiding Principles on Business and Human Rights provide a relevant framework in this respect, particularly through the State duty to protect, the corporate responsibility to respect human rights, and the need to ensure access to effective remedies. In the context of AI, this means that States should regulate corporate conduct in a way that prevents and remedies AI-related human rights harms, while companies should conduct meaningful human rights due diligence throughout the design, development, deployment and use of AI systems.

The monograph has also examined the debate on emerging or new human rights in the digital age. Digital rights should be regarded as an integral part of the contemporary international human rights framework, while the specific catalogue and legal status of emerging digital and AI-related rights remain evolving. These rights and claims reflect concerns about autonomy, identity, participation, accountability, safety and trust in technologically mediated environments. Many AI-related claims may be understood as specifications or extensions of existing rights. Privacy-related protections against harmful uses of AI systems build upon the right to privacy and data protection. Protection against algorithmic bias and unfairness is closely connected to equality and non-discrimination. Accountability for harms caused by AI systems may be linked to effective remedies and due diligence. Algorithmic transparency and explainability may often operate as procedural safeguards supporting due process, access to information and the ability to contest decisions.

At the same time, some proposed AI-related rights reveal the limits of analysing AI-related harms solely through established rights categories. The right not to be subjected to algorithmic manipulation and the right to a human decision and human-to-human interaction are particularly significant in this respect. Algorithmic manipulation does not merely raise questions of data protection or information quality. It concerns the capacity of AI systems to shape preferences, behaviour and decision-making in ways that may

undermine mental integrity, conscious deliberation, autonomy and democratic participation. The right to a human decision and human-to-human interaction similarly raises dignity-related concerns about the replacement of human judgement, empathy and responsibility by automated or AI-mediated processes. These proposed rights should not be understood as absolute. Their scope must be context-sensitive and proportionate. Nevertheless, they indicate that future human rights protection may need to address not only the correction of defective AI outputs, but also the preservation of human agency, dignity and meaningful human involvement in AI-mediated environments.

From a *de lege ferenda* perspective, several priorities may be identified.

First, human rights impact assessments should become a central element of AI governance, particularly for high-risk AI systems used in areas such as law enforcement, migration, employment, education, healthcare, social security and the judiciary. Such assessments should not be limited to technical risk management, but should evaluate the potential impact of AI systems on the full range of human rights. They should be context-sensitive, participatory, iterative and linked to concrete mitigation measures. The development of instruments such as HUDERIA shows that impact assessment can serve as a bridge between general human rights principles and practical AI governance.

Second, procedural safeguards must be strengthened. Individuals affected by AI systems should have access to meaningful information, explanation, contestation and effective remedies. Transparency should not be reduced to formal disclosure or generic technical descriptions. It should enable individuals, courts, regulators and oversight bodies to understand and challenge decisions that affect rights and interests. In this respect, explainability must be adapted to context, purpose and the capacities of the decision-maker or affected person. Where AI systems support high-stakes decisions, human oversight must be meaningful rather than merely formal. Decision-makers who remain responsible for the final decision must understand the role played by the AI system, the limits of its reliability and the reasons why its output should or should not be followed.

Third, the human-centred character of AI governance must be operationalised. AI may function as an empowerment tool by supporting better access to information, improving analytical capacity and assisting decision-making. However, this potential can only be realised where humans are able to form an appropriate understanding of AI systems and their limitations. This

requires training, institutional capacity, appropriate human-AI interaction design, and safeguards against automation bias. Human involvement should not be treated as a symbolic requirement. It must be capable of preserving independent judgement, accountability and the possibility of contestation.

Fourth, institutional coordination should be improved. Human rights bodies, data protection authorities, AI regulators, equality bodies, consumer protection authorities and sectoral regulators should not operate in isolation. AI-related harms often cut across legal fields and institutional mandates. Effective protection therefore requires coordinated supervision, exchange of expertise and clear allocation of responsibility. This is particularly important in relation to systems that affect multiple rights simultaneously, such as predictive policing, automated welfare systems, algorithmic recruitment, biometric identification, credit scoring or AI-supported judicial and administrative decision-making.

Fifth, the role of private actors must be more clearly integrated into AI governance. States should ensure that companies developing or deploying AI systems are subject to adequate legal obligations, including human rights due diligence and access to remedy. At the same time, companies should not be allowed to replace public accountability through voluntary ethics commitments or internal governance mechanisms. The increasing role of private actors in shaping information environments, access to services and decision-making processes makes it necessary to clarify their responsibilities under business and human rights standards and to ensure that public authorities retain effective regulatory and supervisory capacity.

Sixth, international cooperation should be strengthened in order to avoid excessive fragmentation of AI governance. The global nature of AI systems requires greater dialogue between universal and regional organisations, as well as between States with different levels of technological and regulatory capacity. Particular attention should be paid to ensuring that AI governance does not deepen global inequalities or exclude regions with more limited institutional and technological resources from shaping emerging standards. The African and Inter-American perspectives are therefore important not only as regional developments, but also as part of a broader effort to ensure that AI governance remains inclusive, pluralistic and responsive to different social and institutional contexts.

Finally, the debate on new digital and AI-related rights should continue, but with caution. Not every technological risk requires the immediate recognition of a new human right. In many cases, the concerns underlying proposed digital rights may be addressed through the reinterpretation and strengthening of existing guarantees. The recognition of new rights should therefore be considered only where existing rights cannot adequately capture the specific nature of the harm or where a distinct normative interest requires autonomous protection, and where the proposed right can be formulated with sufficient precision to generate identifiable obligations and safeguards. The central question is not whether IHRL should be replaced by a new digital rights framework, but how it can evolve without losing its coherence, universality and normative force.

Overall, the monograph concludes that IHRL is not obsolete in the age of AI. On the contrary, it remains indispensable. Its principles provide the normative foundation for assessing whether AI systems respect human dignity, equality, autonomy, freedom and justice. However, the effectiveness of this framework will depend on its capacity to adapt. This adaptation does not necessarily require a complete reconstruction of IHRL. Rather, it requires the dynamic interpretation of existing rights, the development of effective procedural safeguards, the strengthening of institutional capacity, the integration of human rights into the design, deployment and governance of AI systems, and the cautious articulation of new rights where existing guarantees prove insufficient.

AI therefore does not merely create new regulatory problems. It reveals deeper questions about the future of human rights protection in technologically mediated societies. It challenges IHRL to respond to forms of power that are increasingly data-driven, decentralised, private and opaque. The central task for the coming years will be to ensure that technological innovation remains compatible with the foundational values of human rights law. In this sense, the future of AI governance should not be framed solely as a question of technological control, but as a question of preserving human dignity, agency, accountability and meaningful human freedom in the digital age.

REFERENCE LIST

Monographs

- BAKOŠOVÁ, L.: *Green Ambitions in the Era of Artificial Intelligence*. Praha: Leges, 2024, ISBN: 978-80-7502-788-7, 196 p.
- COPELAND, J.: *Artificial Intelligence: Philosophical Introduction*, New Jersey: Wiley-Blackwell, 1993, ISBN: 978-0-631-18385-3, 336 p.
- DONNELLY, J.: *Universal Human Rights in Theory and Practice*, Second edition, Ithaca: Cornell University Press, 2003, ISBN: 978-0-8014-8776-7, 290 p.
- ELBERT, L.: *Rule of Law a medzinárodné právo*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2015, ISBN 978-80-8152-365-6, 281 p.
- FRANKISH, K., RAMSEY, W. M. (eds.): *The Cambridge Handbook on Artificial Intelligence*, Cambridge: Cambridge University Press, 2014, ISBN: 978-0-521-87142-6, 366 p.
- JANKUV, J.: *Medzinárodné a európske mechanizmy ochrany ľudských práv*. Bratislava: IURA EDITION, 2006, ISBN: 80-8078-096-X, 358 p.
- KEITNER, CH.: *International Law Frameworks (Concepts and Insights)*. Fifth Edition, St. Paul: Foundation Press, 2021, ISBN: 9781647084417, 365 p.
- MAZÁK, J., JÁNOŠIKOVÁ, M. et al.: *The Charter of Fundamental Rights of the European Union in Proceedings before Courts of the Slovak Republic*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2016, 279 p.
- O'FLAHERTY, M.: *Human rights and the UN: practice before the treaty bodies*. Second Edition, Boston: Nijhoff: Brill, Law Specials - Vol. 54, 2021, ISBN: 978-90-04-47928-9, 240 p.
- NILSSON, N. J.: *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press, 2010, ISBN: 978-0-521-12293-1, 707 p.
- STEINER, H. J., ALSTON, P.: *International Human Rights in Context: Law, Politics, Morals*. Second Edition, Oxford: Oxford University Press, 2000, ISBN: 0-19-829849-8, 1497 p.
- ŠTURMA, P.: *Mezinárodní a evropské kontrolní mechanismy v oblasti lidských práv*. 4. vydání, Praha: C. H. Beck, 2024, ISBN: 978-80-7400-961-7, 164 p.

Chapters in monographs

- BALCERZAK, M.: *Implications of the United Nations human rights standards for the development of artificial intelligence*. In: BALCERZAK, M., KAPELAŇSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, pp. 6–44.

- BALCERZAK, M.: *Introduction: AI and International Human Rights Law*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J.: *Artificial Intelligence and International Human Rights Law*, Northampton: Edward Elgar Publishing, 2024, pp. 1–4.
- BENKO, R.: *Právne prostriedky nápravy porušenia základných práv v európskom regióne v komparatívnej perspektíve s osobitným ohľadom na prostriedky nápravy finančnej povahy porušenia práv uznaných Chartou základných práv EÚ v horizontálnych vzťahoch*. In: GIERTL, A. (ed.): *Regionalizmus: Stav, východiská, perspektívy*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2013, pp. 129–141.
- BENKO, R.: *The Impact of Artificial Intelligence on Fundamental Rights and Freedoms and Vice-Versa*. In: KLUČKA, J., BAKOŠOVÁ, L. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Nakladatelství Leges, 2021, pp. 39–50.
- CAROZZA, P. G.: *Human Dignity*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 345–359.
- HANSUNGULE, M.: *The Historical Development of International Human Rights*. In: CHOWDHURY, A. R., BHUIYAN, J. H. (eds.): *An Introduction to International Human Rights Law*, Brill: Nijhoff, 2010, pp. 1–20.
- HEYNS, CH., KILLANDER, M.: *Universality and the Growth of Regional Systems*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 670–697.
- KLUČKA, J.: *General Overview of the Artificial Intelligence and International Law*. In: KLUČKA, J., BAKOŠOVÁ, L., SISÁK, L. (eds.): *Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges*, Praha: Nakladatelství Leges, 2021, pp. 13–24.
- KOTHARI, M.: *From Commission to the Council: Evolution of UN Charter Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 587–620.
- KOVÁCS, P.: *Introductory Thoughts About the Universal Protection of Human Rights*. In: KOVÁCS, P., BERES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, pp. 15–29.
- LAKATOS, I.: *Human Rights Protection in the UN: A General and Institutional Overview (The Former Commission on Human Rights, Human Rights Council, High Commissioner for Human Rights, the International Court of Justice and the Protection of Human Rights)*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, pp. 103–148.
- MARTYNOVA, E.: *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary*. In: BELL, L., GASPAR, W. (eds.): *AI from the Global Majority: Official Outcome of the UN IGF Data and Artificial Intelligence Governance Coalition*. Rio de Janeiro: FGV Direito Rio, 2024, pp. 109–120.

- MORAWSKA, E. H.: *Council of Europe standards and activities related to AI: towards a Framework Convention on AI and human rights?* In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, pp. 25–44.
- POŠIVÁKOVÁ, L.: *Úloha regionalizmu v oblasti ochrany ľudských práv*. In: GIERTL, A. (ed.): *Regionalizmus: Stav, východiská, perspektívy*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach, 2013, pp. 113–128.
- RAJI, I. D. & SMART, A.: *Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing*. In: *FAT*20 Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, 2020, pp. 33–44. Online <https://dl.acm.org/doi/epdf/10.1145/3351095.3372873>
- RAYFUSE, R.: *Public International Law and the Regulation of Emerging Technologies*. In: BROWNSWORD, R., SCOTFORD, E., YEUNG, K. (eds.): *The Oxford Handbook of Law, Regulation and Technology*, Oxford: Oxford University Press, 2017, ISBN: 978-0-19-968083-2, pp. 500–521.
- RODLEY, N. S.: *The Role and Impact of Treaty Bodies*. In: SHELTON, D.: *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press, 2013, pp. 621–648.
- STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J. (eds): *Artificial Intelligence and International Human Rights Law: Developing Standards for a Changing World*. Cheltenham: Edward Elgar Publishing Limited, 2024, pp. 45–67.
- STEIN, A.: *Generative AI and sustainability*. In P. Hacker, P. (ed.): *The Oxford handbook of the foundations and regulation of generative AI*. Oxford: Oxford University Press, 2025. Online: <https://doi.org/10.1093/oxfordhb/9780198940272.001.0001>
- ŠMIGOVÁ, K.: *The History of Human Rights Law: Generations of Human Rights*. In: KOVÁCS, P., BÉRES, N. (eds.): *The Universal Protection of Human Rights*. Miskolc–Budapest: Central European Academic Publishing, 2026, pp. 31–62.
- TYMOFEYEVA, A.: *Jednotlivec jako subjekt mezinárodního práva: Příklad EÚLP*. In: ONDŘEJ, J.: *Postavení jednotlivce v mezinárodním právu: Proměny (vývoj) právního postavení jednotlivce v mezinárodním právu*. Praha: Univerzita Karlova, Právnická fakulta, 2020, pp. 12–23.

Articles

- ALOISI, A., GRAMANO, E.: *Artificial Intelligence Is Watching You at Work. Digital Surveillance, Employee Monitoring, and Regulatory Issues in the EU Context*. In: *Special Issue of Comparative Labor Law & Policy Journal, “Automation, Artificial Intelligence and Labour Protection”*, Vol. 41, No. 1, 2019, pp. 95–121.

- ALSHARIF, M., HNIT, H.: *Digital human rights: Legal debates and emerging foundations under the international bill of human rights*. In: Social Sciences & Humanities Open, Vol. 12, 2025, 102160, pp. 1–10.
- AWAD, E. et al.: *Crowd sourcing Moral Machines*. In: Communications of the ACM, Vol. 63, No. 3, 2020, pp. 48–55.
- BALDINELLI, L.: *Artificial Intelligence in Public Administration: The Importance of AI Impact Assessments between the Framework Convention and the AI Act*. In: Italian Journal of Public Law, Vol. 17, No. 4, 2025, pp. 1375–1389.
- BARONI, M. J. L.: *Fourth Generation Human Rights in View of the Fourth Industrial Revolution*. In: Philosophies, Vol. 9, No. 2, 2024, pp. 39–54.
- BIELIAKOV, K. I., TYKHOMYROV, O. O., RADOVETSKA, L. V., & KOSTENKO, O. V.: *Digital rights in the human rights system*. In: InterEULawEast: Journal for the International and European Law, Economics and Market Integrations, Vol. 10, No. 1, 2023, pp. 183–207.
- BRITO, I.: *The right not to be monitored: in defence of one of the last bastions of the human being against the emerging omniotic surveillance society*. In: UNIO – EU Law Journal, Vol. 11, No. 1, June 2025, pp. 52–87.
- CANCELA-OUTEDA, C.: *The EU's AI act: A framework for collaborative governance*. In: Internet of Things, Vol. 27, 2024, pp. 1–11.
- CAVALCANTE SIEBERT, L., LUPETTI, M. L., AIZENBERG, E. et al.: *Meaningful human control: actionable properties for AI system development*. In: AI Ethics, Vol. 3, 2023, pp. 241–255.
- CUSTERS, B.: *New digital rights: Imagining additional fundamental rights for the digital era*. In: Computer Law & Security Review, Vol. 44, 2022, 105636, pp. 1–13.
- DRACHOVSKÁ, K.: *Umělá inteligence jako nositelka základních práv?* In: Právník, Vol. 4, 2021, pp. 273–284.
- FOLORUNSO, F. A.: *Artificial Intelligence and Human Rights in Africa: Rethinking Protection under the African Charter System*. In: African Journal of Legal Research, Vol. 4, No. 2, 2026, pp. 53–78.
- GRAVETT, W.: *Sentenced by an algorithm — Bias and lack of accuracy in risk-assessment software in the United States criminal justice system*. In: South African Journal of Criminal Justice, Vol. 34, No. 1, pp. 31–54.
- HAENLEIN, M., KAPLAN, A.: *A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence*. In: California Management Review, Vol. 61, No. 4, 2019, pp. 1–10.
- HANNUM, H.: *The Status of the Universal Declaration of Human Rights in National and International Law*. In: Georgia Journal of International and Comparative Law, Vol. 25, 1996, pp. 287–397.

- HENDERSON, P., HU, J., ROMOFF, J., BRUNSKILL, E., JURAFSKY, D., & PINEAU, J.: *Measuring and mitigating the environmental impact of machine learning*. In: *Journal of Machine Learning Research*, Vol. 21, 2020, pp. 1–44.
- HORNEBER, D., LAUMER, S.: *Algorithmic Accountability*. In: *Business & Information Systems Engineering*, Vol. 65, No. 6, 2023, pp. 723–730.
- HUQ, A.: *A Right to A Human Decision*. In: *Virginia Law Review*, Vol. 106, 2020, pp. 611–688.
- CHANG, CH.: *The First Global AI Treaty: Analyzing the Framework Convention on Artificial Intelligence and the EU AI Act*. In: *University of Illinois Law Review Online*, Vol. 86, 2024, pp. 86–99.
- IENCA, M., ANDORNO, R.: *Towards new human rights in the age of neuroscience and neurotechnology*. In: *Society and Policy*, Vol. 13, No. 5, 2017, pp. 1–27.
- JUMA, D.: *Access to the African Court on Human and Peoples' Rights: A Case of the Poacher turned Gamekeeper*. In: *Essex Human Rights Review*, Vol. 4 No. 2, 2007, pp. 1–21.
- KLUČKA, J.: *Úloha a dôležitosť etických pravidiel v systémoch umelej inteligencie*. In: *Právny obzor*, Vol. 108, No. 3, 2025, pp. 240–254.
- KROLL, J. et al.: *Accountable Algorithms*. In: *University of Pennsylvania Law Review*, Vol. 165, No. 3, 2017, pp. 633–705.
- LAZORČÁKOVÁ, T.: *Pristup mimovládnych organizácií a jednotlivcov k africkému systému ľudských práv*. In: *STUDIA IURIDICA Cassoviensia*, Vol. 5, No. 2, 2017, pp. 13–24.
- LECUN, Y., BENGIO, Y., HINTON, G.: *Deep Learning*. In: *Nature*, Vol. 521, 2015, pp. 436–444.
- MACKLEM, P.: *What is International Human Rights Law? Three Applications of a Distributive Account*. In: *McGill Law Journal*, Vol. 52, 2007, pp. 575–604.
- MALGIERI, G., CUSTERS, B.: *Pricing Privacy – the right to know the value of your personal data*. In: *Computer Law and Security Review*, Vol. 34, No. 2, 2018, pp. 289–303.
- MATTHIAS, A.: *The responsibility gap: Ascribing responsibility for the actions of learning automata*. In: *Ethics and Information Technology*, Vol. 6, 2004, pp. 175–183.
- McCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C.: *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, August 31, 1955. In: *AI Magazine*, Vol. 27, No. 4, 2006, pp. 12–14.
- McCRUDDEN, CH.: *Human Dignity and Judicial Interpretation of Human Rights*. In: *The European Journal of International Law*, Vol. 19, No. 4, 2008, pp. 655–724.
- MITTELSTADT, B.: *Principles alone cannot guarantee ethical AI*. In: *Nature Machine Intelligence*, Vol. 1, 2019, pp. 501–507.
- MITTELSTADT, B.D., ALLO, P., TADDEO, M., WACHTER, S., FLORIDI, L.: *The ethics of algorithms: Mapping the debate*. In: *Big Data & Society*, Vol. 3, No. 2, 2016, pp. 1–21.

- MIHAJLOVIĆ, A., ĆORIĆ, V.: *Artificial Intelligence and Discrimination: Strengths and Weaknesses of the Current European Anti-Discrimination Legal Framework*. In: *Regional Law Review*, Vol. 5, 2024, pp. 9–25.
- MORSKA, N.: *The Relationship Between Natural and Digital Human Rights: Problems of Legal Definition*. In: *Philosophy and Governance*, Vol. 18, No. 2, 2026, pp. 1–11.
- MÜLLER, V.C.: *Is it time for robot rights? Moral status in artificial entities*. In: *Ethics and Information Technology*, Vol. 23, 2021, pp. 579–587.
- NAGY, N.: “Humanity’s new frontier”: *Human rights implications of artificial intelligence and new technologies*. In: *Hungarian Journal of Legal Studies*, Vol. 64, No. 2, 2023, pp. 236–267.
- NARDOCCI, C.: *Artificial Intelligence at the Crossroads between the European Union & the Council of Europe: Who Safeguards What & How?* In: *Italian Journal of Public Law*, Vol. 16, Issue 1, 2024, pp. 165–196.
- NOVELLI, C. et al.: *A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities*. In: *European Journal of Risk Regulation*, Vol. 15, 2024, pp. 1–25.
- PAPANASTASIOU, T. N., PAVLIDIS, G.: *The European Court of Human Rights and the Age of Artificial Intelligence: Jurisprudential Foundations for Future Adjudications*. In: *Cyprus Law Review*, Vol. 2–3, 2024, pp. 183–187.
- PARK, P. S., GOLDSTEIN, S., O’GARA, A., CHEN, M., HENDRYCKS, D.: *AI deception: A survey of examples, risks, and potential solutions*. In: *Patterns*, Vol. 5, No. 5, 2024, pp. 1–16.
- PEREPOLKIN, S et al.: *The concept of the fourth generation of human rights: fact or perspective of scientific discourse*. In: *Brazilian Journal of International Law*, Vol. 21, No. 2, 2024, pp. 280–295.
- PETRYSHYN, O. V., HYLIKA, O. S.: *Human Rights in the Digital Age: Challenges, Threats and Prospects*. In: *Journal of the National Academy of Legal Sciences of Ukraine*, Vol. 28, No. 1, 2021, pp. 16–23.
- PRESNO LINERA, M. A., MEUWESE, A.: *Regulating AI from Europe: a joint analysis of the AI Act and the Framework Convention on AI*. In: *The Theory and Practice of Legislation*, Vol. 13, 2025, pp. 1–20.
- PUSZTAHELYI, R.: *Some Reflections on Fundamental Rights Impact Assessment for High-Risk AI Systems (FRIA)*. In: *The Juridical Current Journal*, Vol. 103, No. 4, 2025, pp. 13–35.
- ROIG, A.: *Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR)*. In: *European Journal of Law and Technology*, Vol. 8, No. 3, 2017, pp. 1–17.
- ROTENBERG, M.: *Introductory Note to Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. In: *International Legal Materials*, Vol. 64, 2025, pp. 859–902.

- SZAPPANYOS, M.: *Artificial Intelligence: Is the European Court of Human Rights Prepared?* In: *Acta Humana*, Vol. 93, No. 1, 2023, pp. 93–110.
- SHANY, Y.: *Digital Rights and the Outer Limits of International Human Rights Law*. In: *German Law Journal*, Special Issue: International Law and Digitalization, Vol. 24, 2023, pp. 461–472.
- SONG, L., MA, CH.: *Identifying the fourth generation of human rights in digital era*. In: *International Journal of Legal Discourse*, Vol. 7, No. 1, 2022, pp. 83–111.
- TEO, S.A.: *Human dignity and AI: mapping the contours and utility of human dignity in addressing challenges presented by AI*. In: *Law, Innovation and Technology*, Vol. 15, 2023, pp. 241–279.
- VAN GILS, F., MÜLLER, W., PRÜFER, J.: *Microtargeting, voters' unawareness, and democracy*. In: *The Journal of Law, Economics, and Organization*, Vol. 41, No. 2, 2025, pp. 634–653.
- WACHTER, S., MITTELSTADT, B.D., RUSSELL, CH.: *Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI*. In: *Computer Law & Security Review*, Vol. 41, 105567, 2021, pp. 1–31.

Other academic works

- BOMMASANI, R. et al.: *On the Opportunities and Risks of Foundation Models*, Center for Research on Foundation Models, 2021, 214 p. Online: <https://crfm.stanford.edu/assets/report.pdf>
- CASTELLUCCIA, C., LE MÉTAYER, D.: *Understanding algorithmic decision-making: Opportunities and challenges*. European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA), European Parliament, March 2019, 104 p. Online: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU%282019%29624261_EN.pdf
- ELAM, M., REICH, R.: *Stanford HAI Artificial Intelligence Bill of Rights: A White Paper for Stanford's Institute for Human-Centered Artificial Intelligence*, January 2022. Online: <https://hai.stanford.edu/assets/files/stanford-hai-ai-bill-of-rights.pdf>
- GRAUX, H., GARSTKA, K., MURALI, N., CAVE, J., BOTTERMAN, M.: *Interplay between the AI Act and the EU digital legislative framework*. European Parliament - Policy Department for Transformation, Innovation and Health Directorate-General for Economy, Transformation and Industry, PE 778.575 – October 2025, 108 p. Online: [https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU\(2025\)778575_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU(2025)778575_EN.pdf)
- HOFFMANN, J.: *Technological Determination of AI-Relevant Press and Copyright Law and Generative Content's Relevance for EU Competition Law – The Referral in Case C-250/25, Like Company v. Google Ireland Ltd.*, Max Planck Institute for Innovation and Competition Research Paper No. 25-19, 2025, 22 p. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5411443

- CHENG, M.: *The right to human decision: analyzing policies, ethics, and implementation*. First International Conference on Addressing Socioethical Effects of Artificial Intelligence, 2024, 8 p. Online: <https://openreview.net/pdf?id=VrxvBw0ly>
- NCUBE, C., ORIAKHOGBA, D., RUTENBERG, I., SCHONWETTER, T. (eds.): *Artificial Intelligence and the Law in Africa*, LexisNexis, 2023, 372 p. Online: <https://cipit.strathmore.edu/artificial-intelligence-and-the-law-in-africa/>.
- SARTOR, G., LAGIOIA, F.: *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, STOA | Panel for the Future of Science and Technology, 2020, 100 p. Online: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU\(2020\)641530_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf)
- SHELTON, D.: *An Introduction to the History of International Human Rights Law*. George Washington Law Faculty Publications & Other Works, 1052, 2007, 31 p. Online: https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2045&context=faculty_publications

Internet sources

- ALMADA, M.: *Governing the Black Box of Artificial Intelligence*. In: *SSRN Electronic Journal*, January 2023, pp. 1–26. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4587609
- BABICKÁ, K., GIACOMIN, C.: *Understanding the Scope of the Council of Europe Framework Convention on AI*. *Opinio Juris*, 5 November 2024. Online: <https://opiniojuris.org/2024/11/05/understanding-the-scope-of-the-council-of-europe-framework-convention-on-ai/?utm>
- BÅRDSSEN, A.: *AI and Human Rights at the European Court of Human Rights*, *Verfassungsblog*, 8 December 2025. Online: <https://verfassungsblog.de/ai-and-human-rights-at-the-european-court-of-human-rights/>
- Center for AI & Digital Policy: *Statement of the Center for AI & Digital Policy to the Vice-President of the United Nations Human Rights Council To support the Establishment of the Special Rapporteur for AI and Human Rights*, November 2025. Online: <https://www.caiddp.org/app/download/8585853363/CAIDP-UNHRC-VPRusu-AI-Treaty-09112025.pdf?t=1763220293&utm>
- FERREIRA, J. J., MONTEIRO, S.: *The human-AI relationship in decision-making: AI explanation to support people on justifying their decisions*. In: *Joint Proceedings of the ACM IUI 2021 Workshops*, April 13–17, 2021, 9 p. <https://ceur-ws.org/Vol-2903/IUI-21WS-TESS-7.pdf>
- GOEL, B., NILKANTH, D., SINGH, S.: *AI Governance Beyond 2025: UN Pathways and Implications*. Geneva: Geneva Graduate Institute, 2025, 42 p. Online: https://www.graduateinstitute.ch/sites/internet/files/2025-09/ARP35_Report--2----Shreya-Singh.pdf

- HENDRIX, J.: *Reactions to the Adoption of the UN Global Digital Compact*, TechPolicy. Press, 8 October 2024. Online: <https://www.techpolicy.press/reactions-to-the-adoption-of-the-un-global-digital-compact/>
- KOMAITIS, K.: *Analysis: a brave new reality after the UN's Global Digital Compact*, The Digital Forensic Research Lab, 1 October 2024. Online: <https://dfrlab.org/2024/10/01/analysis-a-brave-new-reality-after-the-uns-global-digital-compact/>
- MELTZER, J., TIELEMANS, A.: *The European Union AI Act: Next steps and issues for building international cooperation*, Global Economy and Development at Brookings, May 2022. Online: <https://www.brookings.edu/research/the-european-union-ai-act-next-steps-and-issues-for-building-international-cooperation-in-ai/>
- MONTGOMERY, K. C., CHESTER, J., KOPP, K.: *Data governance for young people in the commercialized digital environment*, Issue Brief, UNICEF, 2020. Online: UNICEF-Global-Insight-data-governance-commercialization-issue-brief-2020.pdf
- PENAGOS, M., KASSIR, S., VOSLOO, S.: *Policy Brief on National AI Strategies and Children: Reviewing the Landscape and Identifying Windows of Opportunity*, Office of Global Insight and Policy, 2020, 14 p. Online: <https://www.unicef.org/innocenti/media/2516/file/UNICEF-Global-Insight-national-AI-strategy-review-policy-brief.pdf>
- SHANY, Y.: *A Right to a Human-to-Human Interaction*, University of Oxford – Institute for Ethics in AI, 7 November 2024. Online: <https://www.oxford-aiethics.ox.ac.uk/blog/right-human-human-interaction>
- SHANY, Y.: *White Paper - The Need for and Feasibility of an International AI Bill of Human Rights*, Institute for Ethics in AI, University of Oxford, November 2025. Online: <https://aifp.oxford-aiethics.ox.ac.uk/sitefiles/white-paper-professor-yuval-shany.pdf>
- TEO, S.A.: *Artificial Intelligence and Transitional Justice: Framing the Connections*, Impunity Watch 2025, pp. 7–46. Online: <https://www.impunitywatch.org/wp-content/uploads/2025/01/Overview-Paper-Report-Artificial-Intelligence-and-Transitional-Justice-Framing-the-Connections-Impunity-Watch-Jan-2025.pdf>
- Van EST, R., GERRITSEN, J.: *Human rights in the robot age Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality - Expert report written for the Committee on Culture, Science, Education and Media of the Parliamentary Assembly of the Council of Europe (PACE)*. The Hague: Rathenau Instituut, 2017, 58 p. Online: <https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf>
- VILJOEN, F.: *International Human Rights Law: A Short History*. In: UN Chronicle, No. 1 & 2, 2009, pp. 8–13. Online: <https://www.un-ilibrary.org/content/journals/15643913/46/2/3/read>
- WANG, G., ZHAO, J., VAN KLEEK, M., SHADBOLT, N.: *Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children*. Association for Computing Machinery CHI'22 Conference, April 29 – May 5, 2022, New Orleans, LA, USA, 29 p. Online: <https://dl.acm.org/doi/pdf/10.1145/3491102.3502057>

International Treaties

Additional Protocol to the American Convention on Human Rights in the Area of Economic, Social and Cultural Rights: Protocol of San Salvador, San Salvador, 17 November 1988

Additional protocol to Convention 108 regarding supervisory authorities and transborder data flows, ETS No. 181, Strasbourg, 8 November 2001

African Charter on Human and Peoples' Rights, Banjul, 27 June 1981

African Union Convention on Cyber Security and Personal Data Protection, Malabo, 27 June 2014

American Convention on Human Rights: "Pact of San José, Costa Rica", San José, 22 November 1969

Charter of the United Nations, San Francisco, 26 June 1945, 1 UNTS XVI

Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, New York, 10 December 1984

Convention on the Elimination of All Forms of Discrimination against Women. New York, 18 December 1979

Convention on the Rights of the Child, New York, 20 November 1989, UNTS Vol. 1577, p. 3.

Convention on the Rights of Persons with Disabilities, New York, 13 December 2006

Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Strasbourg, 17 May 2024

Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, ETS No. 108, Strasbourg, 28 January 1981

European Convention on Human Rights and Fundamental Freedoms, Rome, 4 October 1950

International Convention for the Protection of All Persons from Enforced Disappearance, New York, 20 December 2006

International Convention on the Elimination of All Forms of Racial Discrimination. New York, 21 December 1965

International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, New York, 18 December 1990

International Covenant on Civil and Political Rights, New York, 16 December 1966

International Covenant on Economic, Social and Cultural Rights, New York, 16 December 1966

Optional Protocol to the International Covenant on Civil and Political Rights, New York, 16 December 1966

Optional Protocol to the Convention on the Elimination of All Forms of Discrimination against Women, New York, 6 October 1999

Optional Protocol to the Convention on the Rights of Persons with Disabilities, New York, 13 December 2006

Optional Protocol to the International Covenant on Economic, Social and Cultural Rights, New York, 10 December 2008

Protocol No. 13 to the Convention for the Protection of Human Rights and Fundamental Freedoms concerning the abolition of the death penalty in all circumstances, Vilnius, 3 May 2002, ETS No. 187

Protocol No. 16 to the Convention on the Protection of Human Rights and Fundamental Freedoms, Strasbourg, 2 October 2013, ETS No. 214

Protocol to the African Charter on Human and Peoples' Rights on the Establishment of an African Court on Human and Peoples' Rights, Ouagadougou, 10 June 1988

Soft Law

Africa Declaration on Artificial Intelligence, Global AI Summit on Artificial Intelligence, 4 April 2025. Online: <https://c4ir.rw/docs/Africa-Declaration-on-Artificial-Intelligence.pdf>

Documents of the United Nations System

Committee on the Rights of the Child: *General Comment No. 12 (2009) on the right of the child to be heard*, CRC/C/GC/12, 20 July 2009

Committee on the Rights of the Child: *General Comment No. 16 (2013) on State obligations regarding the impact of the business sector on children's rights*, CRC/C/GC/16, 17 April 2013

Committee on the Rights of the Child: *General Comment No. 25 (2021) on children's rights in relation to the digital environment*, CRC/C/GC/25, 2 March 2021

ITU: *Child Online Protection Guidelines for children, parents and educators, industry and policy-makers*, 2020. Online: <https://www.itu-cop-guidelines.com/>

OHCHR: *Artificial intelligence must be grounded in human rights, says High Commissioner*, 12 July 2023. Online: <https://www.ohchr.org/en/statements-and-speeches/2023/07/artificial-intelligence-must-be-grounded-human-rights-says-high>

OHCHR: *Business and Human Rights in Technology Project ("B-Tech Project")*: Applying the UN Guiding Principles on Business and Human Rights to Digital Technologies, 25 January 2020. Online: <https://www.ohchr.org/sites/default/files/Documents/Issues/Business/B-Tech/BTechprojectoverview.pdf>

OHCHR: *Mapping report: human rights and new and emerging digital technologies Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/56/45, 20 August 2024

- OHCHR, UNICEF: *Taking a Child Rights-Based Approach to Implementing the UNGPs in the Digital Environment: A contribution to the B-Tech Project*, November 2024. Online: <https://www.unicef.org/childrightsandbusiness/media/1406/file/B-Tech-UNICEF-Briefing.pdf>
- UNESCO: *Recommendation on the Ethics of Artificial Intelligence*, SHS/BIO/PI/2021/1, 23 November 2021.
- UNESCO: *Readiness Assessment Methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence*, SHS/REI/BIO/REC-AIETHICS-TOOL/2023 Rev., 2023
- UNESCO: *Report of the World Commission on the Ethics of Scientific Knowledge and Technology on Robotics Ethics*, SHS/YES/COMEST-10/17/2 REV, September 2017
- UN General Assembly: *High Commissioner for the promotion and protection of all human rights*, A/RES/48/141, 20 December 1993
- UN General Assembly: *Human Rights Council*, A/RES/60/251, 15 March 2006
- UN General Assembly: *International Bill of Human Rights: A Universal Declaration of Human Rights*, A/RES/217(III), 10 December 1948
- UN General Assembly: *Pact for the Future – Annex – Global Digital Compact*, A/RES/79/1, 22 September 2024
- UN General Assembly: *Promotion and protection of human rights in the context of digital technologies*, A/RES/78/213, 19 December 2023
- UN General Assembly: *Promotion and protection of the rights of children*, A/C.3/78/L.19/Rev.1, 8 November 2023
- UN General Assembly: *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, A/RES/73/348, 29 August 2018
- UN General Assembly: *Right to clean, healthy and sustainable environment*, A/RES/76/300, 28 July 2022
- UN General Assembly: *Roadmap for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation: Report of the Secretary General*, A/RES/74/821, 29 May 2020
- UN General Assembly: *Seizing the Opportunities of Safe, Secure, and Trustworthy Artificial Intelligence Systems for Sustainable Development*, A/RES/78/265, 21 March 2024
- UN General Assembly: *Transforming our world: the 2030 Agenda for Sustainable Development*, A/RES/70/1, 21 October 2015
- UN General Assembly: *Terms of reference and modalities for the establishment and functioning of the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance*, A/RES/79/325, 26 August 2025
- UN Human Rights Council: *Impact of disinformation on the enjoyment and realization of human rights: Report of the Human Rights Council Advisory Committee*, A/HRC/RES/61/68, 4 February 2026

- UN Human Rights Council: *Impact of new technologies on the promotion and protection of human rights in the context of assemblies, including peaceful protests* Report of the United Nations High Commissioner for Human Rights, A/HRC/RES/44/24, 24 June 2020
- UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007
- UN Human Rights Council: *Manual of Operations*, August 2008. Online: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.ohchr.org%2Fsites%2Fdefault%2Ffiles%2FDocuments%2FHRBodies%2FSP%2FManual_August_2008.doc&wdOrigin=BROWSELINK
- UN Human Rights Council: *New and emerging digital technologies and human rights*, A/HRC/RES/47/23, 13 July 2021
- UN Human Rights Council: *Non-take-up of rights in the context of social protection: Report of the Special Rapporteur on extreme poverty and human rights*, Olivier De Schutter, A/HRC/RES/50/38, 19 April 2022
- UN Human Rights Council: *Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/47/52, 19 May 2021
- UN Human Rights Council: *Racial and xenophobic discrimination and the use of digital technologies in border and immigration enforcement: Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, E. Tendayi Achiume, A/HRC/RES/48/76, 17 December 2021
- UN Human Rights Council: *Racial discrimination and emerging digital technologies: a human rights analysis: Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, A/HRC/RES/44/57, 18 June 2020
- UN Human Rights Council: *Report of the Special Rapporteur on the right to privacy*, Joseph A. Cannataci, A/HRC/RES/34/60, 24 February 2018
- UN Human Rights Council: *Rights of persons with disabilities: Report of the Special Rapporteur on the rights of persons with disabilities*, A/HRC/RES/49/52, 28 December 2021
- UN Human Rights Council: *Right to privacy in the digital age*, A/HRC/RES/48/4, 7 October 2021
- UN Human Rights Council: *Right to privacy in the digital age*, A/HRC/RES/54/21, 12 October 2023
- UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/51/17, 4 August 2022
- UN Human Rights Council: *The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights*, A/HRC/RES/60/45, 29 August 2025
- UNICEF: *Children's Best Interests in Digital Policy and Practice: From the voices of children, youth and experts around the world*, April 2026. Online: <https://www.unicef.org/>

- innocenti/media/12841/file/UNICEF-Innocenti-Children's-Best-Interests-Digital-Policy-Practice-report-2026.pdf
- UNICEF: *Policy guidance on AI for children: Draft 1.0*, New York: Office of Global Insight and Policy, 2020, 48 p. Online: <https://www.unicef.org/innocenti/media/1326/file/UNICEF-Global-Insight-policy-guidance-AI-children-draft-1.0-2020.pdf>
- UNICEF: *Policy guidance on AI for children: Draft 2.0*, New York: Office of Global Insight and Policy, 2021, 60 p. Online: <https://www.unicef.org/innocenti/media/1341/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf>
- UNICEF Innocenti – Global Office of Research and Foresight: *Guidance on AI and Children: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights 3.0*, Florence: UNICEF Innocenti, 2025, 60 p. Online: <https://www.unicef.org/innocenti/media/11991/file/UNICEF-Innocenti-Guidance-on-AI-and-Children-3-2025.pdf>
- United Nations: *Guidance of the Secretary-General: Human Rights Due Diligence for Digital Technology Use: Guidance of the Secretary-General*, 1 May 2024. Online: [digital-tech-nology-use-guidance-sg-1-en.pdf](https://www.un.org/en/content/common-agenda-report/as-sets/pdf/Common_Agenda_Report_English.pdf)
- United Nations: *Our Common Agenda: Report of the Secretary-General*, New York: United Nations, 2021, Online: https://www.un.org/en/content/common-agenda-report/as-sets/pdf/Common_Agenda_Report_English.pdf
- United Nations: *UN Secretary-General launches AI Advisory Body on risks, opportunities, and international governance of artificial intelligence*, Press release, 2023. Online: https://www.un.org/sites/un2.un.org/files/231025_press-release-aiab.pdf
- United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, eISBN: 9789211067873, 101 p.
- United Nations Inter-Agency Working Group on Artificial Intelligence: *Principles for the Ethical Use of Artificial Intelligence in the United Nations System*, 20 September 2022. Online: https://unsceb.org/sites/default/files/2022-09/Principles%20for%20the%20Ethical%20Use%20of%20AI%20in%20the%20UN%20System_1.pdf
- UN Office of the High Commissioner for Human Rights: *High Commissioner Türk: Human rights are the foundations for AI that benefits humanity: Remarks to AI Impact Summit plenary*, 19 February 2026, Delhi. Online: <https://www.ohchr.org/en/statements-and-speeches/2026/02/high-commissioner-turk-human-rights-are-foundations-ai-benefits>
- UN Office of the High Commissioner for Human Rights, Guiding Principles on Business and Human Rights: *Implementing the United Nations “Protect, Respect and Remedy” Framework*, New York and Geneva, 2011. Online: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinessshr_en.pdf
- UN Secretary-General's High-level Panel on Digital Cooperation: *The Age of Digital Interdependence*, 2019. Online: <https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf>

Documents of the African Union

Africa Declaration on Artificial Intelligence, Global AI Summit on Artificial Intelligence, 4 April 2025. Online: <https://c4ir.rw/docs/Africa-Declaration-on-Artificial-Intelligence.pdf>

ACHPR: *Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa*, ACHPR/Res. 473 (EXT.OS/ XXXI) 2021, 10 March 2021. Online: <https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-human-and-peoples-rights-and-art>

AU: *Agenda 2063: The Africa We Want - Framework Document*, September 2015. Online: https://au.int/sites/default/files/documents/36204-doc-agenda2063_popular_version_en.pdf

AU: *AU Data Policy Framework*, 28 July 2022. Online: <https://au.int/sites/default/files/documents/42078-doc-DATA-POLICY-FRAMEWORKS-2024-ENG-V2.pdf>

AU: *Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity*, July 2024. Online: https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf

AU – STC-CICT: *2019 Sharm El Sheikh Declaration*, AU/STC-CICT-3/MIN//Decl., Sharm El Sheikh, 22–26 October 2019. Online: https://au.int/sites/default/files/decisions/37590-2019_sharm_el_sheikh_declaration_-_stc-cict-3_oct_2019_ver2410-10pm-1rev-2.pdf

Documents of the Council of Europe

Council of Europe – Ad Hoc Committee on Artificial Intelligence: *Feasibility Study*, 17 December 2020, CAHAI(2020)23

Council of Europe – Ad Hoc Committee on Artificial Intelligence: *Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law*, 3 December 2021, CAHAI(2021)09rev

Council of Europe – Commissioner for Human Rights: *The Human Line: Safeguarding Rights and Democracy in the AI Era*, CommHR(2025)57, 20 October 2025. Online: <https://rm.COE.int/commhr-2025-57-report-the-human-line-safeguarding-rights-and-democracy/48802906bd>

Council of Europe – Commissioner for Human Rights: *Unboxing Artificial Intelligence: 10 steps to protect Human Rights*, May 2019, 29 p. Online: <https://rm.COE.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64>

Council of Europe – Committee on Artificial Intelligence: *Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, 7 July 2023, CAI(2023)18

- Council of Europe - Committee on Artificial Intelligence: *Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law (Huderia Methodology)*, CAI(2024)16rev2, 28 November 2024
- Council of Europe – Committee of Ministers: *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, 8 April 2020.
- Council of Europe – Committee of Ministers: *Resolution (99) 50 on the Council of Europe Commissioner for Human Rights*, 7 May 1995
- Council of Europe – Parliamentary Assembly: *Need for democratic governance of artificial intelligence*, Resolution 2341 (2020), 22 October 2020
- Council of Europe – Steering Committee for Human Rights: *CDDH Handbook on Human Rights and Artificial Intelligence*, CDDH(2025)R103 Addendum 3, 12 January 2026. Online: <https://rm.coe.int/steering-committee-for-human-rights-cddh-cddh-handbook-on-human-rights/48802a2f59>
- European Commission for the Efficiency of Justice: *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, Council of Europe: Strasbourg, 3–4 December 2018. Online: <https://rm.COE.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>
- Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, Vilnius, 5 December 2023, Council of Europe Treaty Series – No. 225

Documents of the European Union

- Commission Decision of 24.1.2024 establishing the European Artificial Intelligence Office*, Brussels, 24.1.2024, C(2024) 390 final
- Council Regulation (EC) No 168/2007 of 15 February 2007 establishing a European Union Agency for Fundamental Rights*, OJ L 4M, 8.1.2008
- Declaration on Artificial Intelligence Cooperation*, 10 April 2018. Online: <https://ec.europa.eu/jrc/communities/en/community/digitranscope/document/eu-declaration-cooperation-artificial-intelligence>
- Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council Text with EEA relevance*, OJ L 304.
- Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC*, OJ L, 2024/2853, 18.11.2024.

- European Commission: *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe*, 25 April 2018, COM/2018/237
- European Commission: *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence*, 7 December 2018, COM(2018) 795.
- European Commission: *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Fostering a European Approach to Artificial Intelligence*, Brussels, 21 April 2021, COM(2021) 205 final
- European Data Protection Supervisor: *Towards trustworthy AI in the EU public administration: The EDPS Compass for its new role under the AI Act (2026–2027)*, 2026, 18 p. Online: https://www.edps.europa.eu/system/files/2026-03/26-03-17_edps-compass-under-ai-act-2026-2027_en.pdf
- European Parliament, Council of the European Union, European Commission, Directorate-General for Communications Networks, Content and Technology: *European Declaration on Digital Rights and Principles for the Digital Decade*, OJ C 23, 23.1.2023
- Fundamental Rights Agency: *Assessing High-risk Artificial Intelligence: Fundamental Rights Risks*, 4 December 2025. Online: <https://fra.europa.eu/en/publication/2025/assessing-high-risk-ai>
- Fundamental Rights Agency: *Bias in algorithms - Artificial intelligence and discrimination*, 8 December 2022. Online: <https://fra.europa.eu/en/publication/2022/bias-algorithm>
- Fundamental Rights Agency: *Getting the future right – Artificial intelligence and fundamental rights*, 14 December 2020. Online: <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>
- High-Level Expert Group on Artificial Intelligence: *Ethics Guidelines for Trustworthy AI*, Brussels, 2019, 41 pp. Online: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Charter of Fundamental Rights of the European Union*, 26.10.2012, OJ C 326
- Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive)*, 28.9.2022, COM/2022/496 final
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC*, 4 May 2016, OJ L 119/1

Reference list

Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), OJ L 152

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ L 265

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277

Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on machinery and repealing Directive 2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC, OJ L 165

Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC, OJ L 135

Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L 2023/2854

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 12 July 2024, OJ L 2024/1689

Treaty on European Union (consolidated version), 26.10.2012, OJ C 326, pp. 13–390

Documents of the Organization of American States

American Declaration of the Rights and Duties of Man, Ninth International Conference of American States, Bogotá, 1948

OAS: *Inter-American Guidelines on Data Governance and AI*, 2024, 23 p. Online: <https://www.oas.org/ext/DesktopModules/MVC/OASDnnModules/Views/Item/Download.aspx?type=1&id=1080&lang=1>

Documents of the Organisation for Economic Co-operation and Development

OECD: *Explanatory Memorandum on the Updated OECD Definition of an AI System*. OECD Artificial Intelligence Papers, No. 8, 2024. Online: <https://www.oecd.org/con->

tent/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf

OECD: *OECD Due Diligence Guidance for Responsible AI*, OECD Publishing, Paris, 2026. Online: <https://doi.org/10.1787/41671712-en>

OECD: *OECD Guidelines for Multinational Enterprises on Responsible Business Conduct*, OECD Publishing, Paris, 2023. Online: <https://doi.org/10.1787/81f92357-en>

OECD: *Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 22 May 2019, C/MIN(2019)3/FINAL. Online: [https://one.oecd.org/document/C/MIN\(2019\)3/FINAL/en/pdf](https://one.oecd.org/document/C/MIN(2019)3/FINAL/en/pdf)

OECD: *Revised Recommendation of the Council on Artificial Intelligence*, Council at Ministerial Level, 3 May 2024, C/MIN(2024)16/FINAL. Online: [https://one.oecd.org/document/C/MIN\(2024\)16/FINAL/en/pdf](https://one.oecd.org/document/C/MIN(2024)16/FINAL/en/pdf)

National Legislation and Soft Law

The White House Office of Science and Technology Policy: *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, October 2022. Online: <https://marketingstoragerags.blob.core.windows.net/webfiles/Blueprint-for-an-AI-Bill-of-Rights.pdf>

Decisions of judicial and quasi-judicial organs

African Commission on Human and People's Rights

Social and Economic Rights Action Center (SERAC) and Center for Economic and Social Rights (CESR) / Nigeria - 155/96, African Commission on Human and Peoples' Rights, 27 October 2001

European Court of Human Rights

Big Brother Watch and Others v. the United Kingdom, Applications nos. 58170/13, 62322/14 and 24960/15, Judgment of 25 May 2021

Breyer v. Germany, Application no. 50001/12, Judgment of 30 January 2020

Szabó and Vissy v. Hungary, Application no. 37138/14, Judgment of 6 June 2016

Delfi AS v. Estonia, Application no. 64569/09, Judgment of 16 June 2015

Gaughran v. the United Kingdom, Application no. 45245/15, Judgment of 13 February 2020

Glukhin v. Russia, Application no. 11519/20, Judgment of 4 July 2023

Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary, Application no. 22947/13, Judgment of 2 February 2016

Podchasov v. Russia, Application no. 33696/19, Judgment of 13 February 2024

S. and Marper v. the United Kingdom, Application nos. 30562/04 and 30566/04, Judgment of 4 December 2008

Sanchez v. France, Application no. 45581/15, Judgment of 15 May 2023

Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland, Application no. 931/13, Judgment of 27 June 2017

Roman Zakharov v. Russia, Application no. 47143/06, Judgment of 4 December 2015

Court of Justice of the European Union

Digital Rights Ireland Ltd v Minister for Communications, Joined Cases C-293/12 and C-594/12, Judgment of 8 April 2014

Like Company v. Google Ireland, Case C-250/25, Request for a preliminary ruling, 3 April 2025

Tele2 Sverige AB Post-och telestyrelsen, Joined Cases C-203/15 and C-698/15, Judgment of 21 December 2016

La Quadrature du Net and Others, Joined Cases C-511/18, C-512/18 and C-520/18, Judgment of 6 October 2020

Schrems, Case C-362/14, Judgment of 6 October 2015

Schrems II, Case C-311/18, Judgment of 16 July 2020

SCHUFA Holding (Scoring), Case C-634/21, Judgment of 7 December 2023

Yettel Bulgaria, Case C-806/24, Request for a preliminary ruling, 24 February 2025

Inter-American Court of Human Rights

Interpretation of the American Declaration of the Rights and Duties of Man Within the Framework of Article 64 of the American Convention on Human Rights, Colombia, Advisory Opinion OC-10/89, Inter-American Court of Human Rights, Series A no. 10, IHRL 3429 (IACHR 1989), 14th July 1989

Members of the José Alvear Restrepo Lawyers' Collective v. Colombia, Inter-American Court of Human Rights, Judgment of 18 October 2023, Series C No. 506

ABOUT THE AUTHOR

Mgr. Lucia Bakošová, PhD. (née Bódišová), is an assistant professor and researcher in public international law at the Faculty of Law, Pavol Jozef Šafárik University in Košice, where she completed her doctoral studies in international law in 2020. Her academic work focuses on the interaction between technological progress and public international law, with particular emphasis on international human rights law, international environmental law and the legal regulation of artificial intelligence.

During her academic training and early career, she gained extensive international experience through Erasmus+ study, research, and teaching programmes at several European institutions, including the University of Ljubljana, the University of Málaga, the Mediterranean University of Reggio Calabria, and the University of Naples Federico II. In the course of these exchanges, she lectured on international disaster law and on international legal approaches to artificial intelligence. Her research activities have further included research stays at the International Institute of Humanitarian Law in Sanremo, the European University Institute in Florence, the Alma Mater Research Center for Human-Centered Artificial Intelligence at the University of Bologna, and the University of Lisbon.

She has participated as a principal investigator or researcher in several scientific projects addressing international disaster law, sustainable development, and emerging technologies, including projects focused on artificial intelligence and its ethical and legal implications. Her recent research explores how international human rights law responds to complex and rapidly evolving challenges, particularly in areas where traditional regulatory frameworks are confronted with technological innovation.

She is the author of two monographs and numerous scholarly articles and book chapters on issues of public international law, particularly in the fields of international human rights law, international environmental law, international disaster law, and the legal regulation of artificial intelligence. In recognition of her research activities, she is a two-time recipient of the Rector's Award of Pavol Jozef Šafárik University in Košice and was awarded the Slovak Innovation & Startup Awards (SISA) in 2025 in the category "Scientific Personality of the Košice Universities – UPJŠ Award".

