Deusto Journal of Human Rights
Revista Deusto de Derechos Humanos

BOOK REVIEWS / CRÍTICAS BIBLIOGRÁFICAS

Balcerzak, Michal and Julia Kapelańska-Pręgowska, eds. 2024. *Artificial Intelligence and International Human Rights Law. Developing Standards for a Changing World*. **Cheltenham: Edward Elgar. 347 p.**

Itziar Artíñano Ortiz
https://doi.org/10.18543/djhr.3200
E-published: December 2024

Balcerzak, Michal and Julia Kapelańska-Pręgowska, eds. 2024.
Artificial Intelligence and International Human Rights Law.
***Developing Standards for a Changing World*. Cheltenham:**
Edward Elgar. 347 p.

The book *Artificial Intelligence and International Human Rights Law*, edited by Michał Balcerzak and Julia Kapelańska-Pręgowska, offers a meticulous examination of the interplay between human rights and artificial intelligence (AI) within the framework of international law. The written composition is divided into two parts. The introductory part presents a thorough analysis of the in-force global regulatory frameworks. The final part underscores the importance of conducting specific targeted sectoral evaluations to meticulously examine the ramifications of artificial intelligence across essential domains such as justice, privacy, health, and commerce. The authors and contributors analyze, in detail, the ethical, legal, and political quandaries that artificial intelligence presents across its seventeen chapters. The aim is to formulate strategies that guarantee such technology upholds fundamental rights instead of compromising them. This book's significant contribution to AI ethics and human rights is evident in its comprehensive analysis and the strategies it proposes to ensure AI respects and promotes human rights. Although this purpose aims for a commendable level of effort and ambition, it also reveals weaknesses in integration and coherence that prevent a unified narrative. Nevertheless, their contributions show a notable intellectual rigor and a capacity to illuminate pressing matters within global technological governance.

The book begins by asserting the significant role of global entities, including the United Nations (UN), the Council of Europe, and the European Union (EU), in the governance of artificial intelligence. The first chapter, authored by Michał Balcerzak, focuses on the implications of UN human rights standards in AI governance. Balcerzak critically examines the efforts undertaken by the UN system from 2019 to 2023, highlighting key documents such as the Governing AI for Humanity report, which advocates for ethical and transparent regulation. He emphasizes how existing standards, such as the Guiding Principles on Business and Human Rights, provide a solid normative foundation but

remain insufficient to address challenges like algorithmic bias, mass surveillance, and structural discrimination. This chapter articulates the shortcomings of traditional regulatory frameworks while proposing a combination of ethical principles and practical tools, such as UNESCO's ethical impact assessments, challenging the reader's perspective and understanding of AI governance.

In the second chapter, Elżbieta Hanna Morawska's examination of the Council of Europe's normative advances in AI regulation, particularly the work of the Ad Hoc Committee on Artificial Intelligence (CAHAI) and its evolution into the Committee on Artificial Intelligence (CAI), has practical implications. Her detailed analysis of the fundamental principles underlying the developing Framework Convention, such as transparency, accountability, and data protection, and her call to overcome regulatory fragmentation and ensure non-state actors' inclusion in policymaking provide actionable insights for policymakers and practitioners.

The third chapter, written by the thorough Piotr Staszczyk, is notable for its length and takes an in-depth look at the EU. The Digital Services Regulation 2022/2065, which entered into force on February 17, 2024, has established a more comprehensive framework that classifies AI applications by risk levels and imposes strict requirements for high-impact technologies such as real-time facial recognition. The chapter combines technical analysis with a comparative perspective, exploring tensions between EU regulation and more permissive approaches in the United States and China. The EU's regulatory proposal is presented as a global model for balancing technological innovation with protecting fundamental rights.

In the fourth chapter, Marya Akhtar and Rikke Frank Jørgensen address the impact of automated decision-making systems (ADM) in the public sector. Using practical cases from Denmark, the authors illustrate how ADMs can enhance administrative efficiency, perpetuate discrimination, and infringe on fundamental legal principles. This chapter is precious because it emphasizes algorithmic transparency and ethical impact assessments as essential tools to mitigate ADM-associated risks.

In the fifth chapter, Agnieszka Bień-Kacała examines the impact of Pegasus spyware on human rights and democratic processes, emphasizing its misuse in countries such as Poland and Hungary to surveil political opponents and journalists. Bień-Kacała critiques the inadequacy of existing regulatory frameworks to prevent abuses linked to mass surveillance technologies, proposing measures such as strict prohibitions on their use in politically sensitive contexts. This chapter's

critique of existing regulatory frameworks is a significant contribution to the scholarly debate on AI governance and human rights protection. The chapter combines rigorous empirical analysis with concrete normative proposals, reinforcing its relevance in the contemporary context.

In the sixth chapter, authored by Julia Kapelańska-Pręgowska, Emilia Sarnacka, and Katarzyna Syroka-Marczewska, the interdisciplinary nature of their analysis is highlighted. Their examination of the implications of AI in healthcare, which emphasizes its benefits and risks, combines legal, ethical, and practical perspectives. This comprehensive approach and their emphasis on human oversight as an essential principle for AI governance in health make their chapter stand out. The book's emphasis on human oversight as a practical recommendation for AI governance in health is a significant contribution to the field, as it provides a clear path for policymakers and practitioners to follow.

In the seventh chapter, Joanna Mazur and Zuzanna Choińska analyze the tensions between public security and fundamental rights protection using facial recognition technologies. The authors identify deficiencies in implementing European regulations through case studies such as Clearview AI and advocate for stricter oversight and transparency measures. This chapter is a notable contribution to the surveillance and human rights debate.

In the eighth chapter, Ewa Michałkiewicz-Kądziela examines deepfakes as a growing threat to human rights and democracy. The chapter highlights how these technologies, used to manipulate audiovisual content, can violate privacy, damage reputations, and undermine trust in democratic processes. Michałkiewicz-Kądziela advocates for the global regulation of deepfakes, complemented by educational strategies to strengthen digital literacy.

The ninth chapter, by Ewa Milczarek, addresses the legal challenges posed by AI-generated creativity. It questions traditional concepts of authorship and intellectual property. It proposes innovative solutions, such as contractual models and specific licenses, to ensure that machine-generated works are regulated relatively and effectively. This chapter is a crucial contribution to the debate on intellectual property in the age of AI.

In the tenth chapter, Anne Oloo explores the challenges and opportunities posed by algorithmic media in Africa, analyzing their impact on human rights, democracy, and social structures. The author highlights how AI systems, used to personalize content or combat disinformation, can also exacerbate structural inequalities due to

algorithmic biases and a lack of representativeness in data. A key example is content moderation tools, which often misinterpret or overlook cultural and linguistic nuances, leading to digital exclusion. The author underscores the importance of initiatives like the African Union's Blueprint for AI in Africa, which seeks to foster regional collaboration and harmonize national data protection laws. This holistic approach reflects a commitment to integrating regional particularities into global governance debates, making this chapter essential to understanding how local contexts shape AI implementation and regulation.

In the eleventh chapter, Maria O'Sullivan analyzes the difficulties traditional legal systems face in providing effective remedies to victims of human rights violations caused by AI. The author identifies issues such as algorithmic opacity, which makes it difficult for victims to understand and challenge automated decisions, and the lack of legal mechanisms to address systemic violations created by these technologies. Examples like the SyRI system in the Netherlands or Australia's Robodebt program illustrate how automated tools can perpetuate structural inequalities and infringe on fundamental rights such as privacy and equality before the law when designed without adequate safeguards. O'Sullivan advocates for a comprehensive approach to redress, combining collective measures such as group claims with mandatory ethical impact assessments and greater transparency in system design. This chapter provides a critical perspective on the limitations of existing regulatory frameworks. It emphasizes adapting legal systems to ensure human rights are not compromised in the digital era.

The twelfth chapter, by Joanna Rezmer, focuses on how AI is transforming the labor sector, highlighting its benefits and risks. Rezmer explores how automated technologies reshape recruitment, performance evaluation, and resource management processes, raising serious concerns about worker privacy, decision-making transparency, and perpetuating discriminatory biases. For instance, algorithmic recruitment tools have been shown to reinforce historical patterns of exclusion, disproportionately affecting women and minority groups. The author also examines the risks of job precarity and increasing inequality, noting that while automation can enhance workplace productivity and safety, it can also concentrate economic benefits in the hands of a few. Rezmer calls for adopting international standards to regulate the implementation of AI in the workplace, emphasizing the role of the International Labour Organization in promoting a human-centered approach. This chapter combines rigorous empirical

analysis with concrete normative proposals, making it a critical contribution to the debate on the future of work.

In the thirteenth chapter, Maciej Jerzy Siwicki analyzes the impact of scalper bots in e-commerce, a phenomenon that has gained prominence in the digital economy. Scalper bots are designed to purchase high-demand products and resell them at inflated prices, creating artificial scarcity and harming consumers and businesses. The author examines specific cases, such as mass purchases of event tickets or gaming consoles, to illustrate how these practices negatively impact market fairness and consumer trust. From a regulatory perspective, Siwicki discusses the European legal framework, including the GDPR and the proposed AI Act, highlighting how these instruments aim to regulate bot usage and protect consumer rights. A notable strength of the chapter is its critical analysis of the limitations of existing technical and legal measures and its proposal for a combined approach integrating regulatory solutions, advanced technological tools, and public awareness campaigns.

In the fourteenth chapter, Tomasz Sroka examines how AI is transforming judicial administration, highlighting the benefits this technology offers in terms of efficiency and warning of the risks it poses to fundamental guarantees of the right to a fair trial. Sroka emphasizes that while AI can facilitate administrative tasks and improve predictability in judicial processes, its use in critical decisions raises serious challenges regarding transparency, accountability, and human oversight. Algorithmic opacity, known as the "black box effect", makes it difficult for involved parties to understand and challenge automated decisions, threatening fundamental principles such as equality of arms and the right to adequate defense. The chapter proposes normative safeguards such as mandatory human oversight and algorithmic transparency, stressing that judicial decisions must permanently preserve the central role of human judges. This chapter is particularly relevant in increasing automation, offering a critical reflection on balancing technological innovation with fundamental rights.

In the fifteenth chapter, Agnieszka Szpak addresses the ethical and legal dilemmas associated with lethal autonomous weapons systems, highlighting their growing relevance in international humanitarian law debates. Szpak examines how these technologies, capable of selecting and attacking targets without direct human intervention, pose significant challenges to principles such as the distinction between combatants and civilians, proportionality, and attack precautions. The author discusses debates within the Convention on Certain

Conventional Weapons framework, emphasizing "meaningful human control" as an essential safeguard for ensuring accountability and legality in using these weapons. This chapter combines technical, ethical, and normative analysis, offering a comprehensive view of a topic of growing importance in international security.

In the sixteenth chapter, Lutiana Valadares Fernandes Barbosa and Ana Luísa Zago de Moraes analyze how AI is used in migration processes, highlighting its potential to increase efficiency and the ethical and legal risks it poses. The authors emphasize that using automated systems to determine refugee status can compromise fundamental principles such as due process and the principle of non-refoulement. This chapter stands out for its interdisciplinary approach, combining concrete examples with a solid normative framework, and its proposal of hybrid mechanisms integrating human and technological oversight to ensure fair and transparent decisions.

The seventeenth chapter, authored by Peng Wang and Guannan Qu, explores the implementation of Smart Courts in China. This model has revolutionized judicial administration through AI and big data. The authors describe how this transformation, divided into two main phases —process digitization and the integration of advanced technologies— has enhanced efficiency and transparency in judicial case handling. However, they also highlight these innovations' ethical and regulatory challenges, such as algorithmic opacity, threats to citizens' privacy, and the absence of adequate regulations to standardize application. Despite these challenges, the chapter underscores that while AI provides invaluable support in streamlining judicial systems, it cannot replace human oversight, which is essential to ensuring fairness and public trust. The authors advocate for clear safeguards to balance technological innovation with protecting fundamental rights.

This volume makes a unique and significant contribution to the discourse on AI governance and human rights. It presents a comprehensive research methodology and ethical goals that are commendable. The authors delve into the main challenges that artificial intelligence poses for human rights, covering privacy, justice, employment, and health. However, their reliance on established legal structures and a prevailing normative framework may limit their ability to provide innovative and adaptable solutions. The global landscape is rapidly transforming across numerous industries, driven by advances in artificial intelligence. These developments underscore the inherent deficiencies of international institutions and regulatory frameworks, which need to be revised to effectively address power imbalances and

structural inequities exacerbated by the rise of emerging technologies. The book's failure to offer a comprehensive and inclusive point of view is a significant shortcoming. While it does make superficial allusions to the disparities in technological access, especially in regions such as Africa, it does not comprehensively examine the possibilities for improving these trends. The recognition of pioneering applications of artificial intelligence is evident in fields such as smart agriculture across the African continent. However, we still overlook the ramifications of relying on technologies conceived by external entities. The lack of this point of view reinforces a technocratic narrative that emphasizes technical solutions, overlooking the fundamental structural changes needed to achieve social and technological justice. In addition, the text does not define a feminist perspective sufficiently. While references to the various effects of artificial intelligence on women and marginalized groups are present, they appear fragmented and need a comprehensive, systematic, critical examination. The exclusion of gendered implications and fundamental power dynamics in the discourse around deepfakes and mass surveillance requires scrutiny, as these issues are frequently treated as ancillary rather than central to the analytical framework. The intersectional approach reveals a remarkable mismatch between the prevailing normative narratives expressed in the discourse and the authentic lived experiences of the most marginalized communities.

Nevertheless, this text is fundamental to understanding the intricate relationship between artificial intelligence and human rights. It not only highlights the limitations set by conventional normative approaches but also challenges them, shaping future discussions in the field. The text institutes the international institutions and regulatory frameworks within a historical context characterized by a gradual, rather than an exponential, evolution. In juxtaposition, the emergence of artificial intelligence is simultaneously transforming regulatory structures, leading legislators to get caught up in extensive deliberations that produce minimal substantive results. In light of the rapid advances within technological spheres, it is imperative to conduct a comprehensive analysis of the methodologies used in regulating and implementing artificial intelligence.

In conclusion, this volume significantly contributes to the discourse on AI governance and human rights. Its interdisciplinary approach, sectoral focus, and actionable recommendations provide a robust framework for addressing AI's ethical and legal challenges. The book's value lies in its ability to bridge theoretical principles with practical applications, paving the way for a more equitable and rights-centered

technological future. However, the book could benefit from greater cohesion and a more inclusive perspective, particularly regarding structural inequities and gendered impacts. Despite these limitations, it offers invaluable insights for academics, policymakers, and practitioners navigating the complexities of AI regulation.

Itziar Artíñano Ortiz
Universidad Complutense de Madrid