BART CUSTERS AND EDWARD FOSCH-VILLARONGA (EDS.), *LAW AND ARTIFICIAL INTELLIGENCE: REGULATING AI AND APPLYING AI IN LEGAL PRACTICE**

(INFORMATION TECHNOLOGY AND LAW SERIES, VOL. 35) THE HAGUE: T.M.C. ASSER PRESS AND BERLIN: SPRINGER, 2022 PP. XII+569. ISBN 978-94-6265-522-5

Prof. George Gotsis
Department of the History and Philosophy of Science, National and Kapodistrian University of Athens

This edited volume explores issues regarding the intersection of law and Artificial Intelligence (AI). Law and AI interact in two distinct ways: *law is in position to regulate AI*, yet *AI can be applied in legal practice*. AI embodies a new generation of technologies, primarily characterized for autonomy and self-learning capacity. Concomitantly, AI technologies involve a potential for perennial improvement, *partly irrespective of human intervention*, by engaging in not pre-programmed decision-making processes. Undoubtedly, AI is rapidly transforming the reality we live in, affecting our experience, thus inaugurating a new era of technology that necessitates efficient legal regulation.

The book is divided in five independent parts. Part I consists of introductory chapters on AI technology. Part II and III focus on the need for regulating AI from public law and private law perspectives, respectively. Part II focuses on areas in public law, examining non-discrimination law, labor law, humanitarian law, constitutional law, immigration law, criminal law, and tax law. Part III focuses on fields in private law, among which intellectual property and consumer law. Part IV analyzes applications of AI in legal practice. Part V discusses the future of AI, investigating prospective developments and highlighting implications for legal practice.

The book consists of 27 chapters allocated in accordance

* This paper is an integral part of EU ERASMUS+ Project ETHICS4CHALLENGES (E4C): Innovative Ethics Education for Major Technological and Scientific Challenges.

with the structure referred to above. In the first chapter, the editors provide a panorama of the various themes covered in this book. In ch. 2, Roy De Kleijn, drawing on recent developments on both cognitive science and neural networks, underscores certain similarities and differences between artificial and human intelligence, reviewing the history of their interconnectedness, as well as the evolutionary dynamics of general AI. In ch. 3 Andreas Häuselmann analyzes different definitions of AI, providing a comprehensive discussion of its conceptual dimensions. In so doing, he elaborates on the debate on legal and ethical issues accompanying AI developments, focusing on core sub-disciplines more relevant to law academic community: machine learning, automated reasoning, computer vision, and natural language processing.

The next nine chapters incorporate contributions to public law and AI. In ch. 4, Dolores Morondo-Taramundi examines a set of discriminatory effects in AI-based decision-making. Applying indirect discrimination to algorithm-based decisions is determined by the way in which certain requirements have been developed in EU anti-discrimination law and the EU Court of Justice's case law. One commonality between indirect and algorithmic discrimination is that both are related to *structural patterns of inequality*, typical of hierarchies embedded in social arrangements that reflect the allocation of social privilege. The author concludes by suggesting that hybrid research into algorithmic discrimination may offer technical tools to enhance the performance of anti-discrimination legal reasoning. She claims that the effort to address the shortcomings of anti-discrimination legislation will give further *"impetus to anti-discrimination law and a potential mechanism for transforming inequality structures, rather than only addressing single instances of invidious treatment"* (p. 82). Maria López Belloso in ch. 5 investigates the extent to which normative proposals to regulate the use of AI can reduce the detrimental outcomes of discriminatory biases and protect vulnerable groups. The biases of AI technology attenuate its transformative potential unless it is viewed through the theoretical lenses of diversity and inclusion agendas. The author advances an intersectional approach towards AI-based decision-making, arguing that regulatory frameworks play a critical role in reducing AI gender biases. The author argues that to foster the role of AI in the protection of women's and vulnerable group's rights, we must identify areas where these groups need increased protection so that the deployment of AI technologies will improve these groups' situation (p. 103).

This perspective is specified in ch. 6, in which Edward Fosch-Villaronga and Adam Poulsen explore the precise meaning of diversity and inclusion in an AI context. Not infrequently, discrimination in AI decision processes stems from the AI scientific community's male configuration bias. What is really at stake is *"the amplification of stereotypes, alienation of minority and silent communities, and entrenchment of rigid social expectations in systems"* (p. 127). To help mitigate the risks that missing gender considerations in AI could pose to society, the authors review the inclusion literature, proposing gendering algorithms, more diverse design teams, as well as inclusive policies and regulations. By integrating such considerations, AI systems effectively meet societal needs, respect fundamental rights, and incorporate humanistic values in AI programming.

Elisavet Athanasia Alexiadou in ch. 7 investigates AI in the labor market. Realizing the full potential of AI technologies within employment settings remains a promising task to accomplish from disability rights perspective. AI technologies, if framed in an equity context and in conformity to human rights' principles, are in position to unravel unlawful discriminatory barriers that disabled persons experience in their access to labor markets. In this respect, a disability rights approach to AI is expected to render this type of diversity an integral component of the design, deployment, and assessment of AI. Accordingly, *"a framework for action, encompassing non-discrimination, equity, accountability, transparent and participatory decision-making, which will constitute a roadmap towards building a solid legal basis for the deployment of AI technologies to the actual benefit of persons with disabilities, needs to be developed"* (p. 145).

In ch. 8, Valeria Chiappini Koscina delves into humanitarian law to investigate moral accountability issues in developing lethal autonomous weapon systems. These systems violate the premises of international humanitarian law, primarily because target selections are inherently unpredictable. Under international criminal law, prosecution for criminal responsibility can be attributed primarily in cases of intent and knowledge of a crime. Yet, the likelihood of violating such premises remains insufficient to hold operators accountable for war crimes perpetrated for instance, by killer robots. Commander responsibility seems to be *a suitable mode of liability*, provided that the deployment of such weapons is dependent upon orders reflecting a strictly hierarchical command structure. In a different mode of thinking, Erik Longo in ch. 9 discusses the ways in which the democratic forum is subject to a gradual deterioration, because of the devastating effects of an increasing power of social media platforms. The author convincingly argues that new means of communication enabled through the ongoing digital transformation of the public sphere *do not suffice* to enhance the quality of public deliberation, thus proving ineffective to foster political participation *per se.* This poses a threat to both democracy and liberty, a fact that necessitates regulative initiatives at the EU level. This legal framework *"should treat and protect online behavior differently, better taking into account the 'attention economy' that social media exploits"* (p. 181).

In ch. 10, Clarisse Laupman, Laurianne-Marie Schippers and Marilia Papaléo Gagliardi examine the significance of AI for immigration law. AI has been increasingly employed in several migratory procedures: evaluation of visa applications, processes of determining stay or deportation of immigrants, even in cases of estimating the probability that an asylum seeker is genuinely considered a refugee. However, these AI systems constantly run the risk of undertaking biased decisions, either because of compromised data or because such data were registered by individuals succumbing to social prejudice and biased judgment. Algorithms are subject to biased decision-making due to their programming, a condition severely exacerbated in the case of marginalized individuals, the human rights of which may be thus endangered. If we consider AI as a tool supportive of human agency in migratory processes, then *"it must be ensured that the technology is not being used to undermine human rights, nor is it being used under the discriminatory biases already present in our society"* (p. 202).

In ch. 11, Bart Custers proffers a comprehensive overview of different AI developments

in criminal law. From a substantive criminal law perspective, AI contributes to assessing the effectiveness of sanctions and justice-related practices. From a procedural perspective, AI is employed as a law enforcement technology, for instance in the case of predictive policing, as a cyber agent technology, or by making plausible scenarios when reconstructing criminal acts. Criminal investigation could substantially benefit from AI: this presupposes that discrimination is significantly mitigated in view of enhancing investigation competences and securing fundamental rights.

In the final chapter of the second section Aleksandra Bal explores the benefits from the use of AI in tax law, for instance for detecting VAT fraud in real-time. Focusing on the Polish National Revenue Administration the author assesses the effectiveness of STIR, a tool invented to detect VAT fraud in Poland. Despite the efficacy of this algorithmic process in reducing fraudulent acts, several moral concerns arise regarding STIR's compliance with the principle of proportionality and the right to explanation, both of which are mandated under the EU data protection regulation and human rights legislation. Accordingly, to achieve a satisfactory level of transparency in algorithmic decision-making, it must be ensured that AI produced decisions are communicated and explicated to all affected by them (p. 236).

Chapters 13 to 20 are devoted to issues related to AI applications in private law. In ch. 13, Silvia De Conca discusses AI in the context of liability law: she shows how contractual mechanisms are challenged by AI's technological features. This entails liability gaps for the damaged party, fragmenting the liability among various AI producers, operators, and users. The author suggests that such liability gaps can be addressed through legislative interventions that focus on protecting adversely affected individuals, securing their dignity, rights, and interests. This principle appears to be in alignment with the values underlying EU legislation, thus stimulating the construction of more trustworthy AI systems. In the next chapter, Tycho J. De Graaf and Iris S. Wuisman highlight the fact that the involvement of multiple parties in the operation of AI systems is complicated in cases of a breach of contractual obligations. The authors advocate a legislative proposal by the European Parliament, like that of Dutch civil code that allows the user to escape liability for damage caused by non-high-risk AI systems if the user exercised sufficient care for inspection and maintenance of that system.

In another vein of reasoning, Kostina Prifti, Evert Stamhuis and Klaus Heijne identify an accountability gap in liability law in healthcare. An attempt to address this accountability gap is undertaken by the European Parliament, which advances laws on civil liability for the agents that control AI-systems. Employing a man-machine interaction schema in healthcare, the authors contend that a potential avenue for reform of liability law requires detailed studies, involving an elaborated ethical discourse about how much risk society is ready to bear to make advanced AI-systems available (p. 293). In a same setting, Andrea Bertolini and Shabahang Arian in ch. 16 investigate patient rights in an AI-based healthcare context, focusing on different types of assistance offered by social robots to elderly care. The authors plausibly argue that use of companion robots *does not meet the criteria in national and international legal standards for the right to care*. Insofar as the autonomy, dignity and subjective well-being of older adults is akin

to an ethics of care, AI applications may be deemed as violating the right of the elderly to meaningful care. What is needed is *a deontological approach in accordance with the universality of human dignity* that underlies freedom of self-determination and prevents people from being emotionally manipulated, as prone to an instrumental healthcare ethics.

In ch. 17, Jan Smits and Tijn Borghuis investigate intellectual property law in the case of employing generative AI in music and relevant creative domains. Constructing authorship rights for AI generated works challenges core concepts of intellectual property rights, insofar as autonomous AI systems cannot be attributed the quality of authorship: the attribution of legal personhood to AI systems has not yielded a shared solution. In ch. 18, Hadassah Drukarch and Edward Fosch-Villaronga explore the legal implications of AI-driven boardrooms. Integrating AI-in boardrooms induces major challenges within the field of corporate law, in cases in which *AI systems engage in decision-making independently, operating without constant decision inputs from human directors*. At stake is the issue of attributing responsibility to autonomous systems in cases of detrimental outcomes that AI decisions might generate. The authors posit that *the more autonomous AI systems become, the more decision-making processes shift from human-based to AI-powered entities.*

In ch. 19, Van Duijvenvoorde argues that the interactions between algorithms, the vast amount of available data and online platforms tend to challenge competition law. This entails a re-assessment of the efficiency of competition law, the resurgence of the concept of fairness, and the introduction of harmonization regulation. This mode of thinking culminates in ch. 20, in which Rebecca Owens investigates personalized algorithmic pricing, arguing that competition law embodies the necessary flexibility to resolve hybrid forms of price discrimination. Cooperation with other legislative authorities is deemed necessary to bridge the regulatory gap within AI policies, and ultimately ensure that consumers are sufficiently protected.

The next three chapters are devoted to issues of AI driven legal practice. In ch. 21, Stuart Weinstein explores the use of AI in legal teams. Investigating lawyers' perceptions on the use of AI the author found that while some lawyers believed that machine-learning-induced innovation is in position to transform legal practice, others exhibited little awareness of the opportunities afforded by AI. Apparently, there is a paucity of evidence on whether such tech tools can provide a sustainable competitive advantage to the legal work. This argument is further elaborated in ch. 22, in which Annemarie Drahmann and Anne Meuwese explore AI in lawmaking. They investigate a broad range of AI applications in lawmaking processes, benefiting from the Dutch experience with 'digitization of lawmaking'. AI enables attempts at simplifying the process of developing more uniform legal language, as AI could be used to identify patterns in legislative language (p. 447).

In ch. 23, Antonella Zarra investigates AI developments in the extant legal scholarship. Beyond data protection, highly debated topics remain the regulation of AI applications, ethics, and robots' legal personhood. The last four chapters of the book focus on the future of AI. More specifically, in ch. 24, Bart Verheij recognizes that AI involves practical promises expected

to help legal professionals realize its benefits and invest trust in AI capabilities. Yet, AI has been tremendously evolved that it is by no means evident that AI applications nurture ethics and foster the common good. Accordingly, *constant safeguards are needed for AI to be socially responsible, humane, and ethical*. The author claims that AI ethics is more pertinent to the design and implementation of trustworthy AI systems in alignment with human values, thus suggesting solutions supportive of more humane technologies potentially beneficial to civil society and the public sphere.

Sümeyye Elif Biber and Marianna Capasso in ch. 25 discuss the philosophical and legal dimensions of the right to mental integrity. The chapter focuses on two AI-cognitive human enhancement technologies, brain-computer interface, and intelligent personal assistant. Because of its unprecedented capacity, AI-human enhancement technology runs the risk of not being sufficiently integrated with human cognitive processes, *therefore not serving the goal of mental integrity of individuals* in cases ranging from self-estrangement to distorted perception of human capacities. Thus, we should introduce *absolute* protection to mental integrity in conjunction with mental privacy to protect individuals from any intrusion upon mental states (p. 515).

In ch. 26, Tobias Mahler discusses the societal consequences of Artificial General Intelligence (AGI). AGI is contrasted with narrow definitions of AI. AGI embodies cognitive capabilities beyond those of humans: *such AGI could undergo a perpetual circle of self-improvement, potentially elevating to superintelligence, thus becoming a power factor affecting social evolution*. Albeit such dystopian scenarios remain currently controversial, *the possible emergence of any form of superintelligence necessitates legal regulation* to mitigate long-term detrimental effects. Human oversight and value alignment do not suffice to ensure trustworthy AGI, thus we must assess alternative regulatory mechanisms, particularly in a global context, as EU rules alone may be insufficient for addressing the global challenges of future AI technology (p. 538).

Hin-Yan Liu and Victoria Sobocki in the final chapter examine the prospects of AI policy employing five frames, namely influence, immersion, intensity, integration, and interaction, analyzed through the lenses of potential harm to various stakeholders. The authors identify an overly narrow model of AI which constrains the breadth of policy responses to AI evolution. They in turn suggest an integrated approach centered on distinct types of human relationships with AI, by highlighting the legal implications of involvement in AI-shaped worlds. To address shortcomings of legal protection against detrimental outcomes such as subtle manipulation, the authors suggest policy interventions intended to delineate the problems posed by framing new understandings for AI applications and aligned technologies (p. 558).

In sum, this important book makes a substantial contribution to framing a wide set of issues pertaining to Law and AI research agenda[1]. More specifically, contributions to this

[1] See for instance, J. Gordon, *AI and Law: Ethical, Legal, and Socio-Political Implications,* AI and Society vol. 36 2/2021, pp. 403-404· M. Fink/M. Finck, *Reasoned A(I)Dministration: Explanation requirements in EU Law and the automation of public administration,* European Law Review vol. 47 3/2022, p.p. 376-392· K. Nitta/K. Satoh, *AI applications to the law domain in Japan,* Asian Journal of Law and Society 2022, pp. 471-494· A. Su, *The promise*

volume highlight the spectrum of social consequences of adopting AI in various areas and life domains. More importantly, the authors place an emphasis on the legal framework that should engender AI applications[2] and suggest robust policy interventions to consolidate an effective legal delineation of AI systems. The ultimate rationale for adopting such a regulatory framework remains the need for legal constraint of the dark side of AI[3], the protection and safeguard of human liberty[4] and the promotion of social welfare in both public and private spheres. ❑

and perils of international human rights law for AI governance, Law, Technology and Humans vol. 4 1/2022.

[2] Among the extant literature, please consult, S. Greenstein, *Preserving the rule of law in the era of Artificial Intelligence (AI),* Artificial Intelligence and Law vol. 30 3/2022, pp. 291-323· P. Hacker, *A Legal Framework for AI Training data: From First Principles to the Artificial Intelligence Act",* Law, Innovation and Technology vol. 13 2/2021, pp. 257-301.

[3] On the controversial aspects of AI development, see B.W. Wirtz/J.C. Weyerer/B.J. Sturm, *The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration,* International Journal of Public Administration vol. 43 9/2020, pp. 818-829.

[4] P. Gamez/D. B. Shank/C. Arnold/M. North, *Artificial Virtue: The machine question and perceptions of moral character in artificial moral agents,* AI and Society vol. 35 4/2020, pp. 795-809. See also A. Rachovitsa/N. Johann, *The human rights implications of the use of AI in the digital welfare state: Lessons learned from the Dutch SyRI case,* Human Rights Law Review vol. 22 2/2022.