

ETHICS OF ARTIFICIAL INTELLIGENCE

CASE STUDIES AND OPTIONS FOR ADDRESSING ETHICAL CHALLENGES

Merve Ayşe KIZILASLAN*
ORCID ID: 0000-0001-9654-1423

By Bernd Carsten Stahl, Doris Schroeder, and Rowena Rodrigues. (2023). Cham: SpringerBriefs in Research and Innovation Governance. Doi: 10.1007/978-3-031-17040-9.

Declaration*

In *Ethics of Artificial Intelligence*, authors Bernd Carsten Stahl, Doris Schroeder, and Rowena Rodrigues offer a case-based exploration of the ethical challenges posed by artificial intelligence (AI). Rather than solely engaging in abstract philosophical debate, the authors present a structured, practical analysis regarding how AI interacts with core ethical domains, including discrimination, privacy, manipulation, surveillance capitalism, human dignity, and safety. Contrary to common framing of AI as a purely technical or deterministic force, the book highlights the political, cultural, and social assumptions that underlie AI development and deployment. Through a series of organized chapters, the book succeeds in offering a cohesive examination of what it means to develop AI systems ethically in the 21st century.

The central thesis of the book is that ethical considerations in AI are inseparable from real-world contexts and must be examined through specific, situated examples. Through 21 concise and well-developed cases, the authors bring these issues to life and discuss both the systemic causes of ethical failure and the aspects of possible responses. Their goal is not only to reveal but also to suggest tools such as AI impact assessments and ethics-by-design frameworks.

The authors come from interdisciplinary backgrounds such as philosophy, computer science, law, and public policy. They aim to make a synthesis, reflecting the book's tone, which balances analytical framework with policy relevance. For international relations (IR) and political science scholars, the book provides a precise understanding that AI is not merely a

* Türkiye Cumhuriyeti Cumhurbaşkanlığı İletişim Başkanlığı / Republic of Türkiye Directorate of Communications, Independent Researcher, merve.kizilaslan@iletisim.gov.tr, kizilaslan.merve2224@gmail.com

* This study utilised AI-generated tools. The AI generated examples of what a book review outline should look like. Then, examples from existing literature were examined to illustrate the content of the outline. Finally, at the end of the review, I asked the AI-generated tools to indicate any grammatical errors or sentence corrections needed.



tool of technological innovation, but also a site of governance, ideology, and a field of academic discussion.

The book's primary strength lies in its case study approach. Each chapter follows a clear structure: real-life story, ethical analysis, responses (technical, legal, procedural). This format appears to be ideal for both academic and policy-making uses. Moreover, the authors advocate a pluralist ethical stance, emphasizing the importance of deontology (Kant), consequentialism (Mill), virtue ethics (Aristotle), and care ethics (Held), while recognizing the limitations of a solely Western philosophical perspective.

However, this same structure can sometimes feel repetitive. While the authors are clear about not offering exhaustive philosophical solutions, some readers may find the responses generalized. For instance, tools like “ethics by design” or “AI impact assessments” are well-framed but not discussed enough regarding their real-world adoption or enforcement challenges.

Divided into nine chapters, the book adopts a case-driven and thematically structured approach. It opens with a methodological and philosophical introduction, then moves through specific domains: discrimination, privacy, surveillance capitalism, manipulation, the right to life and liberty, dignity, the UN Sustainable Development Goals, and finally a reflective conclusion.

Each chapter engages a distinct ethical concern, but there is a consistent undercurrent: AI systems do not emerge in a vacuum. They are shaped by historical power conflicts, data inequalities, regulatory vacuums, and socio-political biases. The authors emphasize that ethics must not be reduced to compliance checklists or abstract principles. Rather, it must remain attentive to context, voice, and impact.

The book centralized its ethical concerns by confronting how AI systems can exacerbate structural inequalities, particularly related to gender and race. One crucial example is Amazon's abandoned recruitment tool, which penalized women's resumes due to biased historical data. Similarly, predictive policing tools like COMPAS are critiqued for their lack of transparency and racial bias. The discussions revolve around emphasizing both technical limitations and legal-ethical boundaries (e.g., protected characteristics under human rights law, meaning attributes such as race, gender, religion, or disability, which are protected



against discrimination in legal frameworks to ensure equal treatment and protect human dignity).

Moreover, AI's dependency on large datasets raises questions about privacy, particularly regarding surveillance, genetic data, and biometric information. In that sense, the authors explore China's social credit system and private genomic services like 23andMe, where consent is often shallow and data reuse is unpredictable. Particularly strong is the analysis of "mission creep" and the limitations of the General Data Protection Regulation (GDPR) when applied to AI's evolving capabilities.

Chapters 4 and 5, specifically, mention how AI technologies are weaponized for profit and control. Shoshana Zuboff's surveillance capitalism theory is effectively utilized to explain how companies extract behavioral data to target users. From Clearview AI's biometric scraping to Facebook's microtargeting during elections, the authors underline how opacity, power imbalances, and deceptive interface design erode democratic norms.

The manipulation chapter also highlights how AI is used during user vulnerability, for instance, pushing beauty products during emotionally weak times. These examples problematize the neutrality of algorithmic tools and underscore the ethical costs of optimization-at-all-costs logic.

In the book, the matter of dignity paving the way for philosophically rich and practically urgent discussions. The discussion of how automated decision-making in welfare or healthcare may suppress individuals' voices and recognition is both appropriate and troubling. The proposal for "dignity-sensitive design" and participatory governance serves as a reminder that ethical AI must be not only fair but should consider liberal values while acting in a more humanizing way.

While it appears optimistic by aligning AI with the UN Sustainable Development Goals, it also remains cautious. The authors point out the risk of techno-solutionism, power asymmetries in global AI governance, and the need for reflexivity in design. The suggestion that AI cannot substitute for justice, but may support it if governed wisely, is one of the book's most vital conclusions.

The book's strongest feature is its refusal to separate technology from society. It views ethics not as a simple part of innovation but as integral to design, implementation, and governance.



Its pluralistic approach, from Kantian, utilitarian, virtue ethics, to feminist theories, gathers readers from varied disciplines.

What distinguishes this book is its practical aim. It avoids being merely descriptive by offering governance roadmaps. The authors are aware of the difficulty of embedding ethics into rapidly evolving systems. Hence, they seem to stay cautious about over-relying on principles. Importantly, the book strives to set AI ethics within broader human rights frameworks and social justice discourses, showing how systemic change must accompany technical evolution.

However, one limitation is the weak engagement with non-Western philosophical traditions. While the book acknowledges this gap, more analysis of African, Asian, or Indigenous epistemologies could have enhanced its normative diversity. Furthermore, despite the richness of the case studies, they are mainly drawn from the Global North. Consequently, in order for such a book to address the universal AI ethical approach, it should have broader geopolitical aspects and themes, such as how AI ethics manifest under different state capacities, civil society strengths, and data governance cultures.

Nevertheless, *Ethics of Artificial Intelligence* is a critical read for AI developers, policymakers, related scholars and students. It provides a mirror to current practices and a guide map for future developments. By highlighting how bias, exploitation, and lack of transparency are often embedded in socio-technical infrastructures, the book calls for a radical rethinking of what “ethical AI” really means, not just as a design choice but as a political commitment.

For scholars of ethics, technology, and international affairs, the book serves as both a teaching tool and a research asset. Its clarity makes it suitable for students, while its analytical and case-study-based depth will appeal to academics and policy professionals. The book’s real-world examples, like predictive policing, recommender systems, and biometric identification, ensure it remains understandable and accessible amid the subject’s complexity.

In conclusion, this book succeeds in making AI ethics concrete, relatable, and actionable. Its pluralist methodology, wide-ranging case studies, and commitment to social responsibility make it a valuable contribution to the literature. While it may not satisfy readers seeking in-depth philosophical theorization, it is an exemplary model of applied ethics in the context of emerging technology. As AI continues to shape institutions and everyday life, works like this



are indispensable in guiding ethical and democratic engagement on collective AI moral features.

For anyone concerned with the intersection of technology and society, whether from law, philosophy, international relations, or computer science, this book is a necessary and enlightening read to gain a solid perspective on building AI ethical tasks.

