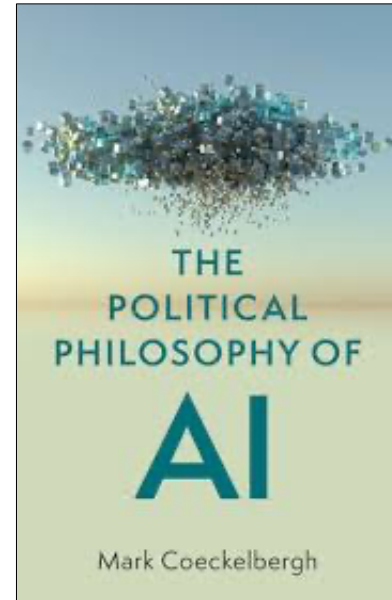


Mark Coeckelbergh, **The Political Philosophy of AI: An Introduction**, Cambridge, UK: Polity Press, 2022, 176 pp., \$22.95 (paperback).

Reviewed by
Chikezie E. Uziegbunam
Rhodes University

Artificial Intelligence (AI) was named the word of the year in 2023 by Collins Dictionary (BBC, 2023). The swift progress in AI and similar emerging technologies has resulted in numerous potentials, although it has also sparked certain concerns. While there is a lot of focus on the ethical issues raised by AI and related technologies like robotics and automation, very little in-depth work approaches the subject from a political-philosophical perspective (p. 3). This is the crux of the book, **The Political Philosophy of AI: An Introduction**, and because it is extensive, capturing its whole essence and divergence in this review is no simple feat. Taking a much-needed political-philosophical turn into the discourse of AI, Mark Coeckelbergh delves into the ethics, social values, and power relations in complex “systems” such as AI. Political philosophy interrogates politics, power, justice, and societal structure to evaluate institutions, authority, and connections within a community. According to Coeckelbergh, in the context of technological advancements like AI and robotics, issues such as those interrogated by political philosophy have become more urgent and relevant (p. 9).



This seven-chapter treatise explores the political philosophy of AI within the context of “political technologies” and “technocapitalism,” emphasizing that AI has both intended and unintended political consequences. AI technologies encompass not only physical tools but also digital platforms, algorithms, and data-driven systems that impact political decision-making, citizen engagement, and the distribution of power. Coeckelbergh explores how AI impacts ideas and values like freedom, social justice, equality, and democracy by analyzing their ethical and political consequences. The essence of the book is encapsulated in its incisive dissection of the political dimensions intrinsic to AI technologies and their consequential effects on pivotal sociopolitical and economic paradigms. The book also argues that the rise of AI necessitates a paradigm shift to include nonhuman entities in our political frameworks, prompting a re-evaluation of power dynamics between humans and AI. The introductory chapter sets the stage for the subsequent chapters by outlining the book’s focus on specific political-philosophical themes indicated above.

Chapter 2 explores the theme of freedom in the context of AI technology and robotics, delving into questions around the political principle of freedom, examining how AI influences decision-making processes, political participations, and individual autonomy, and discussing the implications of algorithmic decision-making and the potential for AI to manipulate and influence human choices. The chapter raises concerns about the impact of digital labor for powerful corporations and explores the concept of robot

slavery in the context of replacing human workers with robots. This chapter examines different interpretations of freedom in political philosophy and their relevance to AI decision-making. It focuses on freedom as political participation in elections and freedom of speech, highlighting challenges such as manipulation, misinformation, and antidemocratic propaganda (see p. 33). Importantly, when talking about the effects of government-run surveillance technologies, the idea of *sousveillance* is brought up as a type of surveillance in which people or groups watch and record activities in response to more traditional forms of surveillance by authorities or institutions. It challenges traditional power dynamics and empowers citizens to hold institutions and authorities accountable for their actions. However, it raises concerns about privacy invasion and the implications of constant monitoring.

Chapter 3 delves into the critical themes of inequality, identity stereotyping, and injustice in the context of AI technology, focusing on bias and discrimination perpetuated by AI systems. The unintended political effects of AI, particularly those enabled by Big Tech companies like Google (p. 40), may exacerbate societal and racialized inequalities. While enhancing efficiency and productivity, AI technologies can also contribute to increasing inequalities in political and economic capital within society. Moreover, the chapter, using philosophical and historical discourses, addresses these concerns, particularly in relation to gender, race, and other marginalized groups.

In chapter 4, the author examines the theme of democracy in the context of AI technology, focusing on the concepts of echo chambers and machine totalitarianism, particularly in the realm of social media and information disorders and how they can impact democratic processes and societal cohesion. For instance, the phenomenon of echo chambers created by personalized algorithms on social media platforms, which reinforce individuals' existing beliefs and limit exposure to diverse perspectives, segregates people into information bubbles. Overall, this hinders political imagination and information diversity and promotes polarization, making consensus and collective action more challenging. Coeckelbergh explores how surveillance capitalism and mass behavior-modification techniques employed by AI technologies can threaten not only personal but also collective political autonomy. This is an extension of the critical work by scholars like Zuboff (2019) on how surveillance capitalism stifles our collective future, undermines citizens' sovereignty, and crushes democratic principles in the digital age.

In chapter 5, the theme of power in the context of AI technology is discoursed, focusing on surveillance and self-disciplining through data, using Marx, Foucault, and other theorists' insights. The chapter explores the implications of AI-powered surveillance on individual autonomy, self-perception, and subjectivity, through data-driven algorithms. The discussion reveals the various ways in which power dynamics manifest in the digital age. First is episodic power—for instance, social media platforms leveraging algorithms to manipulate users' behavior by presenting them with tailored content or advertisements. Second, dispositional power—for instance, empowering students by offering personalized learning experiences and granting access to a vast array of educational resources. Third, systematic power—issues such as data governance underscoring how control and ownership of data can create power differentials between individuals, corporations, and governments, impacting aspects like privacy and autonomy. And finally, constitutive power—social media platforms, for example, influencing how individuals perceive themselves and others through curated content and algorithms that reinforce specific social norms and stereotypes.

Chapter 6 explores the intersection of AI technology with nonhuman entities, specifically questions about the political status of nonhumans, including animals and ecosystems, as well as environmental politics, and posthumanism in the context of AI technology. The author challenges traditional anthropocentric views of politics and ethics by considering the implications of AI for nonhuman beings and the environment. He puts forth a radical proposal for whether AI systems and robots themselves could have political status such as citizenship (pp. 135–137). Drawing on perspectives from posthumanism and transhumanism, readers are invited to reimagine the political and social landscape in a way that includes nonhuman entities.

Taken together, central to this book is the move away from pure technodeterminism to sociopolitical determinism, which advocates for a nuanced understanding of the intricate interplay among technology, entities, ethics, and politics. One fundamental aspect of Coeckelbergh's work is his challenge to the conventional view of technology as a mere means to an end. His perspective underscores the ethical complexities inherent in the development and deployment of AI within political contexts, urging a deeper reflection on the ethical implications of technological advancements. The book is both provocative and broadly illuminating. Its biggest quality lies in advocating for a thoughtful and ethically informed approach to navigating the evolving landscape of technology within political contexts. As AI becomes increasingly integrated into various aspects of daily life, from decision-making processes to social interactions, it is crucial to address issues of bias, discrimination, and power dynamics that may arise from the use of AI systems in different societal contexts.

The book sits powerfully beside not only Zuboff's (2019) work but also Crawford's (2022) *Atlas of AI*, which delves into the social, material, and political dimensions of AI. Written in accessible language, *The Political Philosophy of AI* is a highly useful resource for AI enthusiasts, ethicists, students, researchers, activists, and policymakers interested in comprehending the intricate relationship between AI, politics, and society. As an area for improvement, the author may consider, for the next edition, incorporating more perspectives, particularly from marginalized or underrepresented groups in other parts of the world, as well as concrete practical solutions or recommendations for addressing the challenges identified. Providing actionable insights for policymakers, technologists, and ethicists could enhance the book's impact.

References

- BBC. (2023, November 1). *AI named word of the year by Collins Dictionary*. Retrieved from <https://www.bbc.com/news/entertainment-arts-67271252>
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of Artificial Intelligence*. New Haven, CT: Yale University Press.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. London, UK: Profile.