



## POWER, KNOWLEDGE, AND AI IN POLITICAL DECISION-MAKING

Review of Coeckelbergh, M. (2022) *The Political Philosophy of AI*.  
Cambridge: Polity Press. (186 pages)

Dina Babushkina [d.babushkina@utwente.nl](mailto:d.babushkina@utwente.nl)  
University of Twente

Article type: Review

Review process: Editorial review

This open access article is published with a Creative Commons CC-BY 4.0 license  
<https://creativecommons.org/licenses/by/4.0/>



DOI: [10.59490/jhtr.2023.1.7017](https://doi.org/10.59490/jhtr.2023.1.7017)

ISSN: 2773-2266

Submitted: 15 September 2022 Accepted: 18 September 2023 Published: 10 November 2023

**How to cite (APA):** Babushkina, D (2023). Power, Knowledge, and AI in Political Decision-Making: Review of Coeckelbergh, M (2022). *The Political Philosophy of AI*. *Journal of Human Technology Relations*, 1(1), pp.1-4. <https://doi.org/10.59490/jhtr.2023.1.7017>

Corresponding author: Dina Babushkina

©2023 Dina Babushkina, published by TU Delft OPEN on behalf of the authors

In his book, *The Political Philosophy of AI*, Mark Coeckelbergh shatters illusions concerning the political neutrality of artificial intelligence. His point is not only that AI is relevant to political philosophy, but that it is “political through and through” (Coeckelbergh, 2022, p. 5). What this means is that the role of AI in political processes can no longer be seen merely as curious but a neglected topic for discussion. There is an increasing sense of urgency for philosophers to investigate and evaluate the ways in which this technology affects, challenges, and transforms political processes. Otherwise, Coeckelbergh warns us, we risk finding ourselves—sooner than later—in the absurd world of Kafka’s *The Trial* where those in power are free to pass a decision about our lives without an obligation to provide any rational justification.

What I find most interesting about the book is that it does not merely pave the road to the discussion of different political aspects of AI, but that it does so from a standpoint which is particularly important for the domain of human-technology interactions. The book is motivated by current ethical issues with AI and explores how they play out in the political domain. At the same time, there is a strong interest in the personal rather than purely social dimension of citizenship, which Coeckelbergh approaches from the perspective of decision-making. Essentially, he invites us to consider ways in which AI transforms different aspects of decision making in societal and political contexts: How does AI affect us as citizens in our roles as subjects and objects of political decisions? Are these effects compatible with democratic values? What forms of AI use in the political domain are acceptable and which are not (and how are we to decide)?

The book’s strategy consists of revealing the political nature of AI through two pathways: on the one hand, the book examines the ways AI challenges, transforms, and—to some extent—undermines such core concepts as political freedom, equality, justice, democracy, and power; on the other hand, the book offers an analysis of the ways in which such effects could be approached by different political philosophies. Coeckelbergh examines a good variety of concepts and theories, which makes this book particularly attractive as an entry reading on the topic. Some of these concepts and theories are explored in more depth than some others, but the topic of AI in democratic decision-making clearly stands out. If we were to reconstruct the main argument of the book from the perspective of this topic, it would look something like this.

There is a clear danger that, if left to its own device, AI will threaten the very core of democracy, bringing us to the verge of tech-powered totalitarianism. Different scenarios could unfold, depending on who accumulates power, controlling information through AI: an individual(s), corporation(s) or AI itself (Chapter 4). There are two main risk factors. The first is the fact that reliance on AI technology has a potential to undermine the conditions of our autonomy as political and moral agents (Chapter 1). The second risk factor is the increased inequality and injustice, as well as the tendency to silently normalize both as a consequence of AI use for decision purposes in society (Chapter 2). The key to avoiding AI-steered power imbalances, according to Coeckelbergh, is understanding how power works in relation to technology and knowledge (Chapter 5).

This is insightful: given that AI is a technology of processing and structuring information flows, it is well-equipped to gain control over and manipulate people’s deliberation, decision making, and ultimately choices, be they about private or public matters. In a sense, AI is creating conditions for human cognitive agency to be re-shaped to fit AI’s goals within the societal machinery targeting the control of beliefs (and, often, the way they are generated). This comes with the promise of empowerment through better access to knowledge but may also cost us our agency, i.e., active, and autonomous cognitive agency. From active cognizers, gaining decisional autonomy through reasoning and critical evaluation of information, we are turning into—as Coeckelbergh points out using Couldry & Mejias’ terminology—“inforgs” (Coeckelbergh 2022, p. 106), evolved to be receivers of constant flow of data and adapted to continuous change through nudging. And given the scale of outreach that is open to AI through social media

as well as AI's ability to work on huge data loads, its impact on public opinion can be quick, wide-spread, and devastating.

This leads the reader to conclude together with Coeckelbergh that, ultimately, AI forces us to look deeper into the relationship between knowledge and democratic decision making: How does this relationship relate to the conditions of a democratic deliberation and justification? Can we have political agency without cognitive agency? Does AI carry with it a covert imperative to passively (i.e., uncritically) accept any information targeted to you? What are the potential moral costs of outsourcing part of the decision process to AI?

The book tells the reader that the effects of AI on decision processes are multi-dimensional, and they start with reconditioning decision environments through social media and recommender systems. This means that algorithms actively pre-determine what information you are exposed to as well as the modes of such exposure. One aspect of this is that AI frames the reasons available to us to take into consideration when forming an opinion. As a result, our choice becomes conditioned upon the environment through which the information is filtered rather being dependent on the considerations of what is needed for a rational and well-informed decisions in each case. As a result, we are losing control over our choices without even suspecting that this is happening. From inside the neat AI-shaped information pool that we find ourselves in, we are under the illusion of having access to diverse sources. But AI does more than limit reasons for decision; it narrows the agent's decision horizon by providing more of the same: information is filtered to ensure that it represents the same point of view—depending on what the algorithms calculate to be relevant to the user. This creates tunnel vision, not only significantly limiting the user's exposure to multiplicity of points of view and alternative arguments, but also creating the effect of reinforcement of beliefs, regardless of their truth-status. Apart from increasing the risk of bias, this has potential for the mass-scale manipulation of public beliefs and emotions for political purposes and stirring action among groups of people.

But AI, warns Coeckelbergh, challenges not only the conditions of agency in political decisions. It also changes what it means to be on the receiving end of such decisions. The opacity of AI algorithms threatens explainability of choices that are based on its outcome. This gets in the way of the decision-makers' obligation to give an account of how she arrived at a certain decision in a specific case. The other side of accountability of the decision-maker is a person's right for explanation when a decision is made about her—this is one of the fundamental rights of a citizen, which respects human dignity and keeps the authority of the state in check. The inability of AI system to produce ground for an explanation in terms of reasons and legal norms puts under question the legitimacy of AI as a tool to make decisions about citizens, under the democratic framework. And then again, the very fact that people are forced—often without knowing—to be a constant source of data for AI-augmented political decisions is, in itself, problematic. On the one hand, this creates new forms economic exploitation, i.e., of data-exploiting people for profit. On the other hand, a person, as the object of a political decision, becomes identified with data: AI algorithms determine what counts as relevant information about you in a specific situation. This essentially precludes any possibility for you to control how you are seen by the decision-maker: your beliefs, desires, needs, values, intentions, and (often, even) actions are no longer relevant to you as the object of her decision.

Given this angle on the political nature of AI, the question I would have liked to have seen explored in more detail in the book is whether AI is bringing with it a new paradigm of individual/state relationship, by reshaping the landscape of the individuals' exposure to the state and changing the ways in which she is vulnerable to the abuse of political power. More often than not, we find ourselves bound by an implicit, non-negotiable agreement to supply authorities with data on what we do: How does this affect power recognition mechanisms in society? If power becomes more or less synonymous with the control of data flows, should we normalize the silent acceptance of total personal transparency before an authority with data-

collecting rights? And if the awareness about data-technologies and AI algorithms rises and people object to data-collection, how are we to distribute and limit data-collecting rights?

Coeckelbergh, I think, is right to point that the solution for AI in political domain lies in the question of whether the type of knowledge AI offers is sufficient for decision making in political contexts. This takes us to the right direction if our primary concern is the justifiability of AI-mediated decisions in a democratic context. What seems to be a weak point here, however, is the latent assumption that what AI systems in fact offer is a type of knowledge. If we understand knowledge as a justified true belief, then there are good reasons to think that they do not produce knowledge at all. What seems to be more reasonable to say is that AI processes data to create an information product, but then it all depends on what sort of product this is. It could be information that is not true, it could be information that makes no sense, it could be wrongly interpreted information. In neither case we are dealing with knowledge: What is that, then, that AI produces? Depending on how we answer this question, we need to determine how this AI product relates to the decision processes of an individual. Even more importantly, perhaps, is to understand how AI products should figure in decision processes that satisfy criteria of democratic deliberation. A detailed analysis of the relationship between data, information, and knowledge is needed to solve these problems— an area that is left largely unexplored in the book. This analysis would further help judging whether reliance on AI could lead to a well-justified political decision. However, that task would also require a more detailed look into the criteria of such justification.

To shed light on the way AI mediates our social and political lives, Coeckelbergh's introduces a metaphor he calls "technoperformances". As AI becomes ever more powerful and persuasive, it assumes the role of "choreographer" of humans. The choreographing is total, because it penetrates all areas: how we think, what we feel and need, as well as how we relate to ourselves and our bodies. The analogy does not aim to say that AI becomes a human puppeteer, having its own agenda for the humans, but rather to draw attention to the fact that AI establishes a fundamentally new form of information-based power infrastructure in a society which we cannot escape, but should not lose the right to have an active decision role in. Even though the metaphors of performance and AI as a "choreographer" have an advantage in providing more room—compared, for example, to the mediation theory—for the human role in the way technology shapes the world and ourselves, they remain yet to be developed beyond the limitations of a mere metaphor in order to reveal specific implications of this account of AI. But even so, the take-home message is clear: it is not enough to talk about moral responsibility for AI, you also need to develop and implement social environment and infrastructure for people to be able to exercise their responsible stance towards their—albeit AI-mediated—decisions.

#### Data Access Statement

N/A

#### Contributor Statement

N/A

#### Acknowledgements

N/A

#### References

Coeckelbergh, M. (2022) *The Political Philosophy of AI*. Cambridge: Polity Press.