

Mark Coeckelbergh, *Why AI Undermines Democracy and What to Do About It*

John Wiley & Sons

February 15, 2024

150 pp.

**Mark Coeckelbergh Wants
“Democratic AI” to Build a
Common World**

By Justin Ahn, Stanford University

Since large language models like ChatGPT entered the mainstream, AI has received renewed attention from the general public. To counter techno-pessimism, many commentators repeat the refrain: “AI can be used for good or bad; it just depends on how you use it.” In *Why AI Undermines Democracy and What to Do About It*, Mark Coeckelbergh, Professor of Philosophy of Media and Technology at the University of Vienna, compellingly challenges the notion that AI is politically neutral. He argues that AI will inevitably reshape civic life—for better or worse—and that without proactive policy intervention, it may erode foundational democratic principles.

Coeckelbergh contends that Silicon Valley’s near-total control over AI development disempowers citizens and centralizes power. Like many transformative technologies throughout history, AI risks encouraging more hierarchical, less democratic forms of governance. Because it often operates beyond public oversight, AI can contribute to the ascendancy of capitalism and technocracy over democratic republicanism. (Elon Musk’s post-publication ascent as an unelected advisor to Donald Trump renders Coeckelbergh’s concerns strikingly

prescient.) Moreover, when governments or corporations use algorithms to make decisions affecting individuals, the resulting opacity can deprive citizens of transparency, accountability, and meaningful recourse.

As Coeckelbergh writes:

“We want to steer. We want to steer our own lives and be masters of the new technologies. But quite the opposite is the case. We do not steer but we are steered... At worst, we are the rowers: we are the engines and the fuel of the new digital economy, creating the required data and handing them over to those who use them to enrich themselves at our expense.” (p. 19)

At its most dangerous, AI can pave the way for what Coeckelbergh terms “digital authoritarianism” or even “digital totalitarianism” (p. 5), enabling surveillance, censorship, and the automation of political and bureaucratic decisions that displace human deliberation. Even well-intentioned actors—bureaucrats, designers, or data scientists—may inadvertently create systems that marginalize human agency. More commonly, AI-driven platforms can isolate users, amplify division, and erode the trust and tolerance essential to democratic life.

This concern with communal cohesion is a recurring theme in the book. For Coeckelbergh, AI threatens liberalism not only by challenging liberty and justice, but more deeply by dissolving the relational bonds that sustain democracy. He champions a “thick” conception of

democracy, one rooted not just in majority rule but in participation and dialogue. Drawing on Enlightenment ideals and Habermas's theory of the public sphere, he argues that democracy demands shared world-building:

"the participation in, and co-creation, of a common world" (p. 96).

Yet AI, as currently deployed, tends to fragment that world. Algorithms extract individual preferences and tailor content to isolated users.

"Public opinion becomes a collection of private opinions," Coeckelbergh notes, "rather than a space of genuine exchange and mutual understanding" (p. 46).

Deepfakes, misinformation, and AI-generated "hallucinations" further cloud public discourse, undermining not only trust in institutions but also the very notion of objective truth. On a more intimate level, AI can impair "epistemic agency" by subtly shaping users' choices and beliefs.

"If your political beliefs are manipulated," he asks, "what does your freedom to vote or your right to deliberate and participate in public discussions mean?" (p. 62)

Still, Coeckelbergh remains hopeful that these outcomes are not inevitable. He outlines a vision of "democratic AI," in which public values shape AI design from the beginning. Rather than only regulating harms after deployment, democratic AI would be built to empower citizens, promote

civic education, foster inclusion, and expose users to diverse perspectives. Tools like ChatGPT, for instance, could be reimagined to reduce polarization, counter misinformation, and moderate discourse in democratic spaces.

He also calls for more robust democratic oversight of Big Tech, a revitalized Fourth Estate to hold platforms accountable, and civic education that fosters what he calls "digital humanism." His central conviction is that AI should be governed by democracy, not technocracy—designed to support public deliberation and the common good rather than private profit.

"AI should then help us to realize that richer, relational ideal of democracy: it should help us to build a common world, develop common sense, and work towards the common good" (p. 87).

By rooting AI in shared democratic values, Coeckelbergh offers a timely counterweight to dystopian AI narratives.

At the same time, his analysis opens the door to further questions. In critiquing technocracy, Coeckelbergh highlights a key tension between democracy and expertise. He proposes that citizens define the common good while experts implement it—but readers may wonder how these roles can be better integrated. To his credit, Coeckelbergh gestures toward institutional solutions—a more permanent infrastructure for dialogue between experts and elected representatives—but acknowledges this is an area needing further development.

Likewise, in defending a "thick" democratic ideal, Coeckelbergh is sensitive to the

challenges of balancing civic cohesion with individual rights. His framework casts platforms like TikTok as harmful to democracy due to their isolating effects. Yet he does not endorse outright bans, and instead advocates for thoughtful regulation and design-based interventions to support liberty and mitigate harm. This more nuanced position may not fully resolve the liberty-democracy tension, but it recognizes the complexity of the issue and invites further exploration.

Finally, while Coeckelbergh raises economic alternatives—like platform cooperatives or public ownership of infrastructure—he stops short of endorsing any specific model. This restraint leaves room for debate about whether regulated capitalism can accommodate democratic AI, or whether deeper structural reforms are needed. Rather than prescribing a single solution, Coeckelbergh encourages us to imagine alternatives that preserve both democracy and rights in the digital era.

While he wisely avoids resolving every tension, Coeckelbergh equips readers with a compelling framework—centered on “digital humanism”—to grapple with them. His most enduring contribution is reframing AI’s political impact in terms of deliberation, community, and epistemic agency, rather than simply privacy or automation. He brings rigor to techno-pessimism without falling into fatalism, showing how the structures of digital communication shape knowledge and power.

Perhaps his expansive vision of democratic AI will influence the next generation of algorithmic design. Perhaps new participatory policy models will help create more inclusive systems. For now, *Why AI Undermines Democracy* offers a thoughtful

and provocative starting point. It reminds us that democracy will not survive technological change by default—it must be deliberately reimagined, protected, and practiced.