Journal of Ethics and Emerging Technologies

IEET

*Book Review*

# From Democratic Control to Generative Risk: A Review of Maximilian Kasy's *The Means of Prediction*

**Maximilian Kasy, *The Means of Prediction: How AI Really Works (and Who Benefits)*. (University of Chicago Press, 2025)**

**Juncheng Tao [1,*]**

[1] Center for Logic and Philosophy of Science, KU Leuven, Belgium; juncheng.tao@kuleuven.be
* Correspondence: juncheng.tao@kuleuven.be

**Abstract:** This review examines Maximilian Kasy's *The Means of Prediction* as a critical intervention in debates on AI governance. Kasy argues that the objectives of AI systems are shaped by control over the resources required to build them: data, compute, expertise and energy. By shifting attention from human-machine conflict to conflicts among social groups and interests, the book challenges narratives that present AI development as technically inevitable. This review highlights the strength of Kasy's framework in explaining how AI objectives, prediction and power are connected. It also identifies two limitations. First, the book may understate the importance of technical dimension such as explainability in complex AI systems. Second, its focus on predictive AI leaves open questions about generative and agentic AI, which may reshape the social conditions under which democratic control becomes possible.

**Keywords:** AI governance; Political economy; Democratic control

**Citation:** Tao, Juncheng. 2026. From Democratic Control to Generative Risk: A Review of Maximilian Kasy's *The Means of Prediction*. *Journal of Ethics and Emerging Technologies* 36: 1. https://doi.org/10.55613/jeet.v36i1.237

Received: 30/04/2026
Accepted: 05/05/2026
Published: 05/05/2026



Most debates on AI alignment and AI safety begin from the problem of AI goals or objectives. In technical discussions, this problem is often framed as a failure of optimization (7), captured by terms such as reward hacking, goal misgeneralization, gaming misspecification, or deceptive alignment. Such a framing tends to present the problem as one of misalignment between human intentions and machine optimization. Maximilian Kasy's *The Means of Prediction* shifts the discussion in a different direction. For Kasy, the problem is not only that AI systems may optimize the wrong objectives, but that these objectives are themselves shaped by social, political, and economic power (35).

The book's central logic is simple and powerful. The means of prediction consist of four elements: data, computational infrastructure, technical expertise and energy (83). These resources make AI systems possible, but they are not neutral technical inputs. They are owned, controlled and governed by particular firms, states and social groups. Since the objectives of AI systems are shaped by those who control these resources, democratic governance of AI requires democratic control over the means of prediction. As Kasy argues, understanding AI in its social context requires a shift away from the familiar narrative of human-machine conflict and toward conflicts of values and interests among different members of society (5).

This book makes an important contribution to debates on AI governance by reframing AI as a problem of political economy rather than merely a problem of technical alignment. Its greatest strength lies in showing AI objectives are not simply engineered, but socially produced. At the same time, the book's framework remains largely shaped by the paradigm of predictive AI. This leaves open a further question: whether democratic control over the

means of prediction is sufficient in an era of generative and agentic AI systems that may reshape the very conditions under which democratic control becomes possible.

Kasy's intervention is directed against what he calls the "ideological obfuscation" that dominates much contemporary AI discourse (114). This obfuscation takes several forms. The "existential risk" myth presents AI as a "self-improving force of nature," making its development appear inevitable rather than the result of corporate and political choices (115). The "expert-only" myth suggests that AI is too complex for ordinary people to understand, thereby concentrating authority in the hands of the technological industry and its affiliated experts (116). The "political inevitability" myth invokes geopolitical rivalry to justify rapid and largely unchecked AI development (116). Against these myths, Kasy insists that AI is not fate. It is a matter of collective human decision.

The book's argument can be organized around three connected themes: predictive AI and decision problems, the means of prediction, and democratic control. These themes link prediction to power. What AI predicts, how it optimizes, and whose interests its decision serve are never purely technical questions. Kasy begins by showing that the success of AI is not simply a technical achievement, but a form of predictive success embedded in decision-making. AI systems are often conceptualized as "rational agents" that choose actions by "maximizing some notion of a measurable reward or minimizing some notion of a measurable loss" (23). Yet Kasy's main concern is not with the specific algorithms used to solve automated decision-making problems. Rather, he is interested in the social problem contained within the decision problem itself (35).

This point matters because many social decisions are now being reformulated as predictive problems. AI systems may be used to predict who will be admitted to universities, who will be hired or fired by companies, or who will be targeted in warfare (31). Once such decisions are translated into automated prediction and optimization, the crucial question becomes what resources make prediction possible and who controls them.

Kasy identifies four main resources as the means of prediction: data, compute, expertise and energy. Each is both a technical input and a site of political-economic power. Data can be enclosed by firms that train models on public texts, online interactions and user behavior (85). Compute depends on costly chips, cloud infrastructure, and server farms, creating tendencies toward monopoly and geopolitical competition (90). Expertise gives technical workers bargaining power, since successful model-building requires not only formal knowledge but also practical experience and trial-and-error (92). Energy links AI development to material and environmental constraints (95). Together, these resources explain why control over AI objectives is concentrated in the hands of firms and states that can mobilize them at scale.

Kasy's conclusion follows from this analysis: the control of AI objectives is determined by control over the resources required to build AI. His proposed response is democratic control over the means of prediction. He defends this both instrumentally and intrinsically (112). Instrumentally, democratic control could help align AI systems with broader social welfare rather than with the interests of a small number of firms or states. Intrinsically, democracy is necessary because AI systems increasingly shape decisions that affect society as a whole. Those affected by AI systems should therefore have a say in how these systems are designed, deployed, and governed.

Although the book does not provide a fully detailed institutional blueprint, Kasy identifies several possible agents and mechanisms of change (99). Workers matter because AI depends on highly paid technical experts, low-paid data-labeling workers, and algorithmically managed workers such as delivery riders or drivers. Consumers may also constrain firms, especially when individual choices are amplified by media and public debate. Legal and institutional arrangements are equally important, antitrust law, privacy law, intellectual property law and antidiscrimination law all help determine who controls AI systems and

whose interests they serve (108). In this sense, democratic control does not mean only public participation in abstract debates. It also requires actual rules that redistribute power over data, infrastructure, labor and decision-making.

The book covers many more topics than can be discussed here, including fairness, automation, value alignment, etc. Its main achievement, however, is not that it offers a new technical account of AI. Rather, it changes the political imagination through which AI is understood. Kasy explains AI in a way that is accessible without being simplistic. For example, he compares utility in economics with reward in AI, showing why reward is not merely descriptive but prescriptive, and therefore socially controversial (25). His narrative style is concise and clear. As he suggests, there is no such thing as a person who is simply "not the type" to understand AI (8). In this respect, the book itself functions as a public intervention: it challenges the idea that AI governance should be left to experts alone.

However, Kasy's social perspective also has limitations. Its clarity sometimes comes at the cost of oversimplification. One important question concerns the extent to which AI decision problems can be understood primarily as social and political problems, and the extent to which technical issues of system behavior remain independently important. Kasy is right to resist narratives that reduce AI debates to a conflict between humans and machines. Yet technical opacity means that machine behavior can still become a relatively autonomous source of risk, even when the social objective is known.

This issue is especially clear in relation to explainability (186). Knowing what an algorithm is supposed to maximize is not always enough to understand what it actually does or how reliable its predictions are. In complex AI systems, especially large neural networks, learned behavior may diverge from the formal objective. A system trained to predict exam performance or maximize engagement may rely on hidden proxies, spurious correlations, or emergent patterns that are not visible from the stated decision problem alone. Transparency is therefore not the same as explainability. Disclosing an AI system's objective, data, code, or even internal weights does not necessarily make the system intelligible. A neural network may be fully open access, yet its millions or billions of parameters may still fail to provide a humanly meaningful explanation of why a particular decision was made.

A further limitation is that Kasy's framework remains largely shaped by the paradigm of predictive AI, although he briefly discusses chatbots in a short subsection (53). His core questions—who defines the objective, who controls the means of prediction, and whose interests are served—are highly illuminating for systems that operate within already defined decision problems. Yet current AI development is increasingly centered on generative and agentic systems, which raise different problems for democratic governance.

Kasy asks who controls the means of prediction. In the context of generative AI, we must also ask who and how affects the conditions of democratic control. Generative AI systems do not merely optimize within existing democratic institutions. They may transform the conditions under which democratic control becomes possible in the first place. By producing persuasive text, synthetic media, personalized misinformation, automated political communication, and artificial social interaction, these systems can reshape citizens' informational and social environments before democratic control even begins. The risk is therefore not only that AI systems may be governed undemocratically, but that they may weaken the agents, publics, and institutions capable of governing them democratically.

This suggests that democratic control over the means of prediction must be supplemented by democratic protection of the means of prediction. Predictive systems may distort decisions within existing institutions; generative systems may distort the communicative environment through which citizens form preferences, deliberate with others, and hold institutions accountable. In the age of generative and agentic AI, the preconditions of democratic control are themselves under pressure. Democratic governance of AI must

therefore defend not only fair objectives and accountable institutions, but also the social and epistemic conditions that make democracy possible.

These limitations do not undermine the value of Kasy's book. Rather, they show where his framework needs to be extended. The book's central argument remains powerful: AI should not be treated as an autonomous destiny, but as a field of collective choice. Its greatest strength lies in shifting attention from technical failures alone to the social forces that shape AI objectives, resources, and outcomes. For readers interested in AI safety, political economy, or technology governance, Kasy's book offers a clear, accessible, and provocative framework. Yet the rise of generative AI also shows that the task ahead is not only to democratize control over the means of prediction. It is also to protect the democratic capacities through which societies remain able to decide, collectively and meaningfully, what AI should be for.

**Use of AI tools:** ChatGPT5 was used only for language polishing and formatting. The ideas, arguments and final responsibility for the content remain the author's own.

**Funding:** No external funding was received for this book review.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

Kasy, Maximilian. (2025) The Means of Prediction: How AI Really Works (and Who Benefits), University of Chicago Press.