Journal of Ethics and Emerging Technologies

IEET

Book Review

Review of *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*

Arvind Narayanan & Sayash Kapoor, *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference* (Princeton, Princeton University Press, 2024)

Simon Boulter [1]*

[1] Independent Researcher
* Correspondence: simonboulter1991@gmail.com, ORCID: 0009-0003-9545-7838

Abstract: Reading Arvind Narayanan and Sayash Kapoor's *AI Snake Oil* within broader debates concerning technology, language, and political economy, this review positions AI as a ubiquitous, authoritative, and yet ambiguous category which simultaneously obscures and amplifies the wide range of technologies, practices, and concerns it connotes. Through close engagement with the book's critique of emerging AI technologies and the institutional production of technological hype, the review highlights its attempt to reclaim public discourse on artificial intelligence from both commercial mystification and apocalyptic speculation. At the same time, it suggests that important questions concerning AI's relationship to the production of knowledge remain insufficiently explored.

Keywords: AI; Ethics and Governance; Democracy; Artificial Intelligence; regulation

Citation: Boulter, Simon. 2026. Review of *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*. *Journal of Ethics and Emerging Technologies* 36: 2. https://doi.org/10.55613/jeet.v36i2.245

Received: 01/06/2026
Accepted: 10/06/2026
Published: 01/07/2026

Publisher's Note: IEET stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

In his book *Plastic Words: The Tyranny of a Modular Language*, Uwe Pörksen identified the emergence of a new form of language which has surreptitiously seeped into everyday speech, subtly undermining the ways in which people both communicate and think. Plastic words, as Pörksen calls them, are instantly familiar and at the same time mysteriously opaque – terms borrowed from vernacular language, imported into a scientific field, and returned to everyday use with new, and somewhat ambiguous, application and authority. Common instances might be 'identity,' 'management,' 'progress,' or 'resource': seemingly ordinary words whose plasticity resides in their broad, imprecise accessibility, which, rather than diminishing their effectiveness, confers on them both quasi-technical authority and universal currency. Their application, Pörksen concludes, not only displaces more precise language in favour of scientific-sounding yet unspecific terms but also becomes less evident the more pervasive it grows. In short, plastic words are words that have lost their roots. They are both everywhere and nowhere.

In the decades since Pörksen's analyses (first published in 1988), the ambiguous hold these terms exert, which is no less powerful for its ineffability, has only strengthened. They form a distinct class, their vague atmosphere of objectivity performing distinct roles across the languages of politics, business, and academia, and numbering now perhaps many more than the thirty or forty initially identified by Pörksen. Indeed, in their recent book *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*, Arvind Narayanan and Sayash Kapoor effectively nominate another word that has rapidly risen to prominence in the contemporary lexicon of plastic words: 'AI' – a term whose imprecision and pervasiveness has profound consequences for the ways in which the wide range of technologies it denotes are understood, deployed, and governed.

The extraordinary breadth of these technologies is outlined in the book's introduction, which likens using the umbrella term 'AI' to designate the field of loosely related technologies which fall beneath it as using, for instance, the collective noun 'vehicle' to refer to all forms of transportation: from cars and buses to bicycles and spacecraft. It goes without saying that conversations in such a world would be challenging. And yet, for Narayanan and Kapoor, it is not an unapt analogy for the process of trying to conduct meaningful discussion about the function and future of AI. To make matters worse, in this interpretive vacuum, wilful or negligent commercial, institutional, and intellectual practices of distortion, omission, and amplification abound. Financial actors and fraudsters capitalise on consumer uncertainty. Researchers and universities, increasingly incentivised to drive 'engagement', risk overstating or sensationalising findings. News organisations dealing with commercial pressures and constrained resources can quickly generate clicks by echoing either positive or negative industry hype. In short, Narayanan and Kapoor's book attempts to comprehend, expose, and in part redress this trend.

The conceptual challenge of achieving this feat is made no less tricky by influential recent publications in the field. In 2021, Henry Kissinger, Eric Schmidt, and Daniel Huttenlocher published *The Age of AI*, a book whose shortcomings are illustrative of much of the discussion around AI. Outlandish quotes such as the following confer upon AI a quasi-mystical intelligence which exceeds the grasp and limits of human understanding or imagination:

> The advent of AI obliges us to confront whether there is a form of logic that humans have not achieved or cannot achieve, exploring aspects of reality we have never known and may never directly know (Huttenlocher, Schmidt, & Kissinger, p. 16).

Narayanan and Kapoor's work is adept at exposing the overblown nature of such rhetoric and, at the same time, remains resistant to the equally pervasive countermovement of condemning AI in equally apocalyptic tones, which, in turn, often serves merely to intensify the hysteria surrounding it. Indeed, to the contrary, they highlight how, compared to biological systems like humans, we possess a great deal of insight into how AI functions and can therefore confidently assert both its shortcomings and practical benefits. They point to a more complex network which allows AI hype or mania to take root:

> If we lack a scientific understanding of some aspects of AI, it's because we've invested too little in researching it compared to the investment in building AI. And when we lack an understanding of a specific AI product, it's usually because the company has closed it off to scrutiny (pp. 252-253).

In this vein, the authors attempt to wrest control of the AI narratives from the vested interests that seek to muddy the waters for their own gain. "We are not okay with leaving the future of AI up to the people currently in charge", they write (p. 289).

In spite of the book's title, what follows is a largely balanced account of the hype, misinformation, and misunderstanding that surrounds AI. Readers' reception of its overall findings and balance will no doubt be guided in no small part by their political views, particularly vis-à-vis the state of contemporary capitalism, as much as their stance on AI technologies themselves. This is perhaps because the history and present conversation surrounding AI reflects broader cultural dynamics captured by Ted Chiang's perspective that "fears about technology are fears about capitalism" (Wheeler, 2023). In this light, *AI Snake Oil* self-avowedly advocates a 'middle ground', allowing its

analyses to emerge through close reading of real-world case studies, rather than pre-ordained polemical positions. The resulting critique of big tech and the powerful institutional ecosystems that surround AI proves to be all the more compelling precisely for this reason. At the same time, nuanced opinion and technical expertise is also encouraged to enter the debate with the ambition of informing readers exactly how AI technologies work and, perhaps more importantly, why they often don't.

The authors develop their study by outlining an important distinction between the two dominant forms of AI: predictive and generative. Predictive AI works by making predictions about the future in order to guide decision-making in the present: which candidate will perform most effectively if hired for a particular job, or how many crimes are likely to occur in an area at a specific time. In contrast, generative AI works by developing 'content' in seconds: chatbots create often-realistic answers to human prompts, and image generators produce photorealistic images matching any description. Much of the authors' scorn is directed at predictive AI technologies. It is here they identify a large proportion of the 'AI snake oil'—"AI that does not and cannot work as advertised"—being peddled (p. 2). Predictive AI, they convincingly elaborate in chapters two and three through a range of real-world examples, "not only does not work today, but will likely never work because of the inherent difficulties in predicting human behavior" (p. 3). Chapter four shifts the focus to generative AI. Here the authors are more sympathetic, noting that they are both enthusiastic users of generative AI and demonstrating how and why they "think it can be useful for most knowledge workers" (p. 99). However, they also dedicate large sections of the chapter uncovering what they call 'automated bullshit': "speech that is intended to persuade without regard for the truth" (p. 139). AI chatbots can be "shockingly good at sounding convincing on any conceivable topic", they write, although much of the 'knowledge' it communicates merely derives from regurgitated internet text or incomplete training materials (p. 139). In short, generative AI is optimised to sound natural, not to possess or verify the material it communicates. Nor can it wholly be divorced from its application for malicious practices such as the generation of deepfakes or unscrupulous business practices. In response, the authors conclude that better research and regulation, rather than avoidance, is the best course for combating these issues.

While the recent wave of generative-AI-based chatbots represents the first time this form of AI technology has been *useful* to a large number of people (not just businesses and governments), the technologies most people are confronting are rather a consequence of the steady accumulation of fifty years of progress in the field. Chapter five unpacks this dichotomy in an attempt to illustrate why so much contemporary commentary surrounding AI has become explicitly existential, with non-experts often turning to themes and ideas borrowed from science fiction and Hollywood to better understand the implications of these 'new' technologies. While it is "hard to predict the nature and capabilities" of yet-to-be-invented AI technologies, the authors are emphatic that they see civilisational vulnerabilities coming from well-known and highly predictable sources rather than AI: the fragility of democracy, weapons of mass destruction, climate change, public health, global financial infrastructure (pp. 177-178). In portraying AI technologies as omnipotent, critics overstate its capabilities and underemphasise its limitations, playing into the hands of tech companies who would prefer less scrutiny and distracting from more nuanced research into the benefits or harms the technologies may bring to bear on the broader civilisational challenges to come.

Chapter 6 addresses one of these vulnerabilities in detail, focusing on attempts to moderate social media. Content moderation, the reader learns, is often performed inadequately by AI tools which struggle to understand context or nuance. Moderation itself, however, is made much trickier by the very nature of social media platforms, which

set themselves up as vehicles of entertainment, tools for social connection, and global public squares all in the same movement. The inherent issues in safeguarding these spaces are a fundamental part of their multifaceted and malleable design and application. Moreover, the closer content becomes to violating moderation policies the more likely it is to be promoted and engaged with – that is, recommendation algorithms actively amplify content that is likely to be in some way problematic. This has created a curious tug-of-war within the AI ecosystem: recommendation algorithms amplify potentially troubling content, while moderation algorithms try to detect and suppress it. Much of people's online existence, the authors conclude, operates in the grey areas which exist between these two forces.

And yet, while this existence appears to be changing rapidly in stride with technological advances and emerging technologies, Narayanan and Kapoor are direct in stating that the underlying paradigms controlling and guiding AI are stable – most declarations of new products revolutionising the field are likely to be hype, and the underlying institutional and commercial forces are well established. This does not mean, they suggest, that the work of adapting or restricting institutions and practices in their wake should be neglected, but rather that the work of policy makers should be holistic, patient, and long-term rather than reactive, impulsive, or merely performative. In this way, the text engages successfully with the political and commercial questions raised by AI. For policy makers or those interested in the broader cultural and commercial ramifications of these technologies, it will serve as a useful guide to the state of play. While new stories about AI continually flood the media, those interested in cultivating a deeper perspective will find this book a valuable guide for both navigating the territory and locating possible or alternative paths.

While the authors have clearly thought critically about this issue themselves, a recurring shortcoming of the study is its failure to investigate the role of AI in discouraging precisely the critical, long-term, or unstructured thinking necessary in tackling the problems it generates, particularly so among the young. While the book is adept at illustrating how AI systems are, or are not, operating, the more significant, far-reaching, and yet subterranean question remains unaddressed: not what AI tools are doing, but what they are saying to and about humanity. If the prevalence of plastic words symbolises a desire to shortcut or bypass deep or collective reflection, it is perhaps telling that the quintessential plastic word of the present epoch should reflect precisely this – an *artificial* form of intelligence. Roberto Bazlen once said that in his time intelligence had ended up in the hands of the stupid. It is high time to turn the imagination toward rethinking the ramifications of leaving so much of contemporary human intelligence at the disposal of a tool which, while it may be 'intelligent', is also *essentially stupid* in the sense that it *cannot even think*.

References

Huttenlocher, Daniel., Schmidt, Eric. & Kissinger, Henry. (2021). *The Age of AI: And Our Human Future*, New York: Little Brown and Company.

Narayanan, Arvind. & Kapoor, Sayash. (2024). *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*, Princeton: Princeton University Press.

Pörksen, Uwe. (1995). *Plastic Words: The Tyranny of a Modular Language*, trans. Jutta Mason & David Cayley. Pennsylvania: The Pennsylvania Sate University Press.

Wheeler, Grant. (2023). "'Autocomplete on Steroids': Ted Chiang Addresses Phenomenon of AI at Granfalloon Festival." *Indiana Daily Student*, 9th June 2023. www.idsnews.com/article/2023/06/buskirk-chumley-theater-event-ted-chiang-talk-2023-granfalloon.