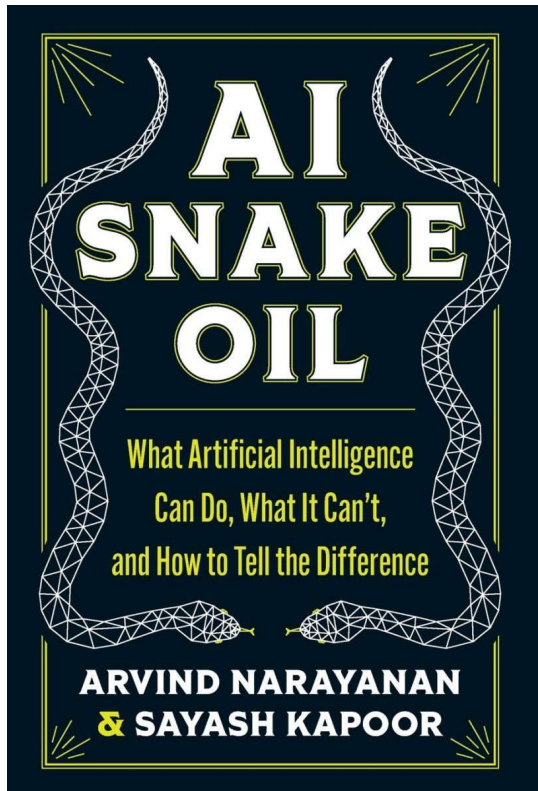


‘AI Snake Oil’ dives into technology’s capabilities

Posted: Sunday, January 5, 2025

“AI Snake Oil: What Artificial Intelligence Can Do, What It Can’t, and How to Tell the Difference”
by Arvind Narayanan and Sayash Kapoor. Princeton, N.J.: Princeton University Press, 2024, 360
pages, \$24.95 (hardcover).



“AI snake oil is AI that does not and cannot work as advertised,” Arvind Narayanan and Sayash Kapoor explain in the introductory chapter of “AI Snake Oil: What Artificial Intelligence Can Do, What It Can’t, and How to Tell the Difference,” their recently released treatise that needs to be read by anyone enamored with the seemingly infinite possibilities currently associated with Artificial Intelligence. “Since AI refers to a vast array of technologies and applications, most people cannot yet fluently distinguish which types of AI are actually capable of functioning as promised and which are simply snake oil.”

“This is a major societal problem: we need to be able to separate the wheat from the chaff if we are to make full use of what AI has to offer while protecting ourselves from its possible harms, harms which in many cases are already occurring,” the authors continue. “A deeper understanding of AI will both satisfy your scientific curiosity and translate into practical ideas on how to use – and when not to use – AI in your life and career.”

Given all the hype that has accompanied the meteoric rise of AI over the last couple of years, I found “AI Snake Oil” to be a much-needed antidote to the undeserved adulation and ever-expanding mythology surrounding applications such as ChatGPT and similar programs. In academia, we are always infatuated with the latest technological development and its usually over-rated potential to revolutionize everything we do. If I had a dollar for every time I’ve heard, “But AI is different; it really is a gamechanger,” I could probably retire tomorrow. And if you work in higher education, I’m sure you remember hearing the same thing about online education. All the faculty teaching face-to-face courses were supposed to be out of a job by now.

“Let’s be clear – AI is no threat to education, any more than the introduction of the calculator was,” the authors assert. “With the right oversight, it can be a valuable learning tool.”

Yes, there are some amazing things that AI has taken beyond the realm of possibility. As Narayanan and Kapoor readily admit, generative AI is making significant inroads into the fabric of both education and business. The applications of this next generation of computing are virtually endless. When it comes to predictive AI, however, the prognosis is significantly less inspiring. Or as the authors put it, “... we will make the argument that predictive AI not only does not work today but will likely never work, because of the inherent difficulties in predicting human behavior.”

Apparently, predictive AI only works accurately and consistently when it is trained on data that is representative of the context in which it will be used. If the environment and circumstances evolve and diverge from the environment and circumstances which served to train the AI, then problems will inevitably arise – problems that can have life and death consequences.

Interestingly, included in the examples provided regarding the problems with predictive AI is United Health: “In one extreme case, U.S. health insurance company UnitedHealth forced employees to agree with AI decisions even when the decisions were incorrect, under the threat of being fired if they disagreed with the AI too many times. It was later found that over 90 percent of the decisions made by AI were incorrect.”

Structurally, the book consists of eight deceptively complex chapters: “Introduction,” “How Predictive AI Goes Wrong,” “Why Can’t AI Predict the Future?” “The Long Road to Generative AI,” “Is Advanced AI an Existential Threat?” “Why Can’t AI Fix Social Media?” “Why Do Myths about AI Persist?” and “Where Do We Go from Here?” One of the more extensively researched efforts I’ve had occasion to read lately – with 38 pages of source notes at the conclusion of the main narrative – I found the prose surprisingly easy to decipher. Both advanced practitioners as well as the more casual reader with an interest in AI should be able to navigate the concepts articulated in what is, in the final analysis, a cautionary tale.

A Professor of Computer Science at Princeton University, Narayanan is Director of the Center for Information Technology Policy. His previous books include “Bitcoin and Cryptocurrency Technologies” with Joseph Bonneau and Edward Felten, and “Fairness and Machine Learning,” with Solon Barocas and Moritz Hardt. Kapoor is a PhD candidate in Computer Science at Princeton University who previously worked as a software engineer at Facebook, where he helped create AI for content moderation.

“Some new AI technologies will hopefully one day come to be seen as mundane,” Narayanan and Kapoor conclude. “Today, self-driving cars often make the news for accidents and fatalities. But safe automated driving is ultimately a solvable problem, although one whose difficulty has repeatedly been underestimated. The bigger challenge for society might be the massive labor displacement that the technology will cause if it becomes widespread – millions of people drive trucks, taxis, or rideshare vehicles. Still, if the safety problem is solved and the necessary social and political adjustments are made, we may one day take self-driving cars for granted, like we do elevators today.”

A recurring theme in the book is the somewhat dubious – yet remarkably pervasive - claim by first adopters that AI will solve previously unsolvable problems and make our lives unimaginably better on all fronts. The authors argue persuasively and convincingly that this declaration has been grossly exaggerated and completely oversold. Just like the idea that if you aren’t starting to integrate AI into every aspect of your personal and professional life, you are somehow “missing out” and will very shortly be left behind as the rest of society continues its ever-ascending trek to high-tech paradise.

As Julia Angwin, author of “Dragnet Nation: A Quest for Privacy, Security, and Freedom in a World of Relentless Surveillance,” notes, “If you are only going to read one book on AI, make it this one. Narayanan and Kapoor make simple what everyone else is trying to make complex. Reading it makes you realize you are not crazy, the AI discourse is crazy, and they have given you a roadmap to clarity.”

After making my way through this extraordinarily enlightening manuscript, I could not agree more. Highly recommended.

Reviewed by Aaron W. Hughey, University Distinguished Professor, Department of Counseling and Student Affairs, WKU.