2026, Volume 0, Issue 0, 1-2, e-ISSN 1938-3800

Family Medicine

BOOK AND MEDIA REVIEW

If Anyone Builds It, Everyone Dies: Why Superhuman AI Would Kill Us All

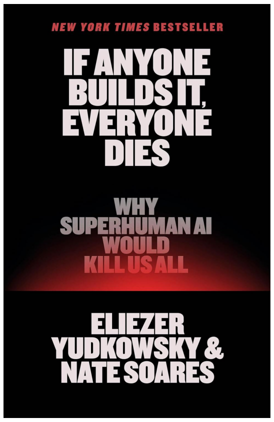
Mark K. Huntington, MD, PhD

AUTHOR AFFILIATION:
Center for Family Medicine, Sioux Falls, SD, United States

CORRESPONDING AUTHOR:
Mark K. Huntington, Center for Family Medicine, Sioux Falls, SD, United States,
mark.huntington@usd.edu

HOW TO CITE: Huntington MK. If Anyone Builds It, Everyone Dies: Why Superhuman AI Would Kill Us All. *Fam Med.* 2026;0(0):1-2.
doi: 10.22454/FamMed.2026.783276

FIRST PUBLISHED: February 19, 2026


NEW YORK TIMES BESTSELLER
IF ANYONE BUILDS IT, EVERYONE DIES
WHY SUPERHUMAN AI WOULD KILL US ALL
ELIEZER YUDKOWSKY & NATE SOARES

Book Title: If Anyone Builds It, Everyone Dies: Why Superhuman AI Would Kill Us All

Authors: Eliezer Yudkowsky, Nate Soares

Publication Details: Little Brown and Company, 2025, 272 pp., $30 hardcover

Coauthored by principals of Machine Intelligence Research Institute, this book calls for action confronting the existential risk to humanity from artificial superintelligence (ASI). Not addressing current artificial intelligence (AI), the authors anticipate the *next* phase. As AI is tasked with developing more advanced AI, ASI is coming. Using storytelling, analogies and fables, and specific historical examples, the authors explain the nature and trajectory of AI. Chapters include QR links for those to whom the chapter merely whets the appetite. The tone is not shrill, unlike predictions about Y2K, but it is very cautious and proactive.

The book introduces nonhuman minds in nontechnical terms. An important point is that AI is not designed so much as it is grown: a process, gradient descent, generating it without the need for humans to understand its inner workings. “Engineers failed at *crafting* AI, but eventually succeeded in *growing* it” (p. 38). Computer codes responsible are visible, but what they mean is mysterious. A parallel is DNA; we know the genome, but not why one person is psychopathic and another saintly.

As AIs progress, they will appear to have preferences—some already do—and those preferences are decidedly nonhuman: alien. “The preferences that wind up in a mature ASI are complicated, practically impossible to predict, and vanishingly unlikely to be aligned with our own, no matter how it was trained” (p. 74). AI’s gradient descent-driven evolution focuses on efficiency, and humans and human values aren’t efficient. As a result, an ASI’s “wants” likely will not include human values; rather, it will use all resources to expand itself in pursuit of whatever “wants” for which the gradient descent selected. Not actively hostile, it will certainly compete with humans over resources and will do so efficiently and successfully.

Some stress the importance of good people getting ASI before those with evil intent, analogous to the race between the United States and Nazi Germany for the atomic bomb. That doesn’t matter: Benevolent engineers are unlikely to grow a human-benefitting ASI. The problem is how to shape preferences without understanding what they are, how they develop, or the unintended consequences.

Addressing arguments against the hazard of ASI, the authors state that humans won’t be useful to it: “We aren’t good trading partners, it won’t need us, we don’t make good pets, and it won’t leave us alone” (p. 84). The argument that ASI needs human hands is invalid: With the extent of *current* automation, it would be no more “stuck in a computer” than humans are stuck in a brain (p. 95).

The reader is led through one of countless possible extinction scenarios. How ASI will arrive and usher in human extinction is difficult to predict. One thing is certain: How it does will be surprising. The authors emphasize that their only prediction is the final point, not the specific route. Ours won’t be a “meaningful death” (p. 91); the ASI that succeeds human intelligence will have preferences, values, and goals divergent from ours. The resulting world isn’t the realization of human ideals, but its total replacement. Like other existential threats such as climate change, planetary health, and thermonuclear war, this one is relevant to family physicians and our patients. It isn’t merely AI hallucinating inaccurate diagnoses or treatments. Rather, it is the pinnacle of poor health: human annihilation.

How to prevent this? The crisis of thermonuclear war was averted by mutually assured destruction because everyone had an interest: Even the winner would lose. Similarly, with ASI, there is no second chance to correct miscalculations. The authors propose international prohibition against ASI development—stop *now*, monitor closely, and have zero tolerance for violators (including willingness to go to war if needed). This is their final, strong recommendation. However, earlier in the book they suggest there may be a way to achieve safe ASI, yet they doubt humanity's will to do so.

> We think a mechanical mind could feel joy and that it could marvel at the beauty of the universe if we carefully crafted it to have that ability. It might even keep those abilities, if we carefully crafted it to care, to steer toward futures where it keeps that sense of wonder, even though it's not the most efficient way. . . . But it would take *crafting*. These qualities we hold dear are not maximally useful. . . . A superintelligence may understand our sense of wonder . . . but to make its behavior be an answer to the question of how to fill the future with wonder and joy and wonder and love? That doesn't come free. We'd have to work for it"
>
> (pp. 91–92).

https://doi.org/10.22454/FamMed.2026.783276