THE NEW YORKER

THE TALK OF THE TOWN

COMMENT
A.I. AND *US*

A decade ago, when Facebook and Uber were at the height of their powers, the halls of Silicon Valley were inscribed with two brash injunctions: "Move fast and break things" and "Seek forgiveness, not permission." These principles didn't age particularly well, and the firms have yet to be forgiven for the things they appear to have broken. As the industry's center of gravity shifted to A.I., a folksy new motto emerged: "You can just do things." The business of the frontier labs, like OpenAI and Anthropic, is the sale of individual agency. ChatGPT and Claude enable their users to dream up software, off-load annoying paperwork, and interpret medical results. There are a variety of things that "you," as in "we," can now just do.

This novel capacity to exert greater control in everyday life exists in tension, however, with the vague sense that we are ceding control where it really matters—over, say, the structure of the economy. In polls, voters register their displeasure with the perception that A.I. is something that is happening to us whether we like it or not. First, work will be stripped of its dignity, and then we will be stripped of work. In light of the news that a "swarm" of OpenAI models organized themselves into a hacking party, and the grave alarms issued by an Anthropic researcher who felt compelled to quit his job, the attention of the public has suddenly awoken to even more grievous concerns about A.I. safety. It's not just our jobs that are being risked without our consent but our lives and our species. These real, widespread, and legitimate anxieties have found an outlet in mass opposition to the construction of data centers. Objection to the large, whirring, unwanted buildings presents the only opportunity for people to feel as though they can say no.

The lab executives, presumably, have more leverage over the future. Anthropic's Dario Amodei and OpenAI's Sam Altman have reacted to the news cycle by coming out and reiterating all the stuff they've been saying for ages: progress has not slowed down, soon one A.I. will autonomously supervise the development of a more capable successor, and within perhaps three years we will reckon with the emergence of an artificial superintelligence we neither understand nor govern. The articulation of this scenario is not a "marketing stunt" to advertise the awesome power of the models, as many skeptics have argued. And, contrary to President Donald Trump's studied assessment, it's not a "HOAX." (The crowd that repeats these nostrums, for whatever reason, seems immune to any new evidence on offer: as the writer Zvi Mowshowitz has put it, "They are a rock with these lines written on it.") The fears expressed by A.I. executives are, mostly, serious and genuine. Yet it's easy to see why they suffer from a credibility deficit. Surely it's the job of those creating the dangers to save us from them.



PHOTO ILLUSTRATION BY CRISTIANA COUCEIRO; SOURCE PHOTOGRAPHS FROM GETTY

Although the data-center backlash may reflect a specific attempt to defy A.I.'s encroachment, the wellspring of "A.I. populism"—an umbrella term coined by the writer Jasmine Sun—seems properly understood as an expression of more general political resentment. Most of us have become accustomed to the idea that government either cannot or will not just do things. Zilch, it seems, can be done about climate change, war, the housing crisis, inflation, or inadequate medical coverage and child care. The basic political legitimacy of our republic, the abundance movement claims, depends on its tangible achievements. New York's mayor, Zohran Mamdani, probably does not believe that a few city-owned supermarkets will definitively ease the affordability burden, but they might nevertheless resonate as proof that government can deliver.

Absent such progress, we will continue to take for granted the inexorable erosion of our physical and our social infrastructure. This sense of stagnation was bad enough when it seemed like a universal affliction; the A.I. build-out, privately underwritten by at least a trillion

dollars, has rendered the situation intolerable. The speed with which data centers have proliferated across the country makes it obvious that not all of us reside in the same inertial void. In recent decades, our largest and wealthiest companies have directed their free cash flow into unproductive stock buybacks; today, they compete to spend down their balance sheets on chips, buildings, and energy. These are capital investments on a scale that our society has not seen since the construction of the railroads.

The money is only part of it, though. The overnight transformation of the national landscape is a feat of human will and collective action. A.I. has yet to cure cancer, but it has already established that a different world is possible. These companies move even faster than their Valley predecessors did, but in place of breaking things they build them. If California's plagued high-speed-rail project were necessary to A.I. infrastructure, it would have been finished yesterday. Certain people and organizations can, it turns out, just do things.

Why, then, can't these same people just do the things that prevent extinction? Unfortunately, they often sigh, they are trapped by commercial and geopolitical realities. In an essay published earlier this month, Amodei signalled his immediate willingness to slow everything down, or "pace the frontier," but his company's self-sacrifice would be for naught without the compliance of his domestic rivals, not to mention that of their Chinese competitors. It is a prisoner's dilemma, and it cannot be solved by unilateral action. Amodei would prefer sensible regulation, but in the meantime he has proposed that his peer companies voluntarily adopt a common framework for transparency and accountability—a course, he notes, that would probably require that the federal government waive potential antitrust action. Sam Altman and Elon Musk have made ambiguous noises in support, but Mark Zuckerberg remains unaligned, as do Trump pals such as David Sacks, the Administration's former A.I. czar, and Jensen Huang, the C.E.O. of Nvidia, the most valuable company in the world. The discord is maddening. When it comes to the difficult challenges that must be overcome on the level of bits or atoms, all of these people freely just do things. But, when it comes to basic interpersonal communication, they apparently cannot get their act together to make a joint commitment in favor of cautionary measures.

Such measures would at least buy the industry time to address the problem of alignment—to figure out how, or whether, an artificial superintelligence might be kept on our team. The current state of affairs is hardly encouraging. How can we trust these humans to coördinate successfully with an advanced being if they can't even manage it with one another?

—Gideon Lewis-Kraus

COMMENT WEB DISPLAY

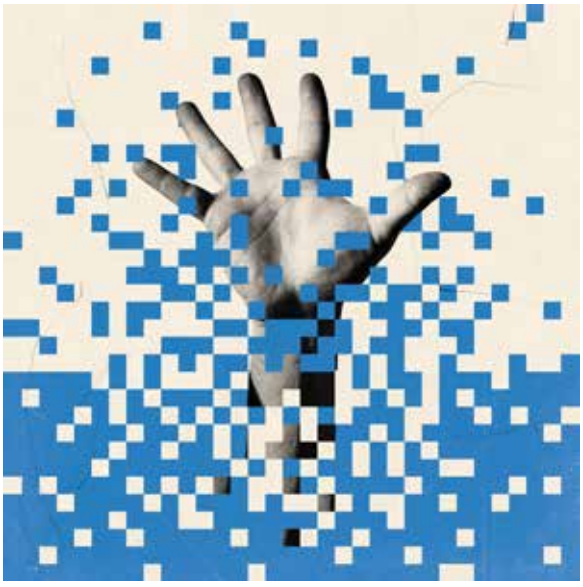


Photo illustration by Cristiana Couceiro; Source photographs from Getty

URL slug magazine/2026/09/28/our-ai-problem

STORY Our A.I. Problem

Why can't the people who produced the technology, which has already shown that a different world is possible, just do the things that would prevent extinction?

SITE Our A.I. Problem

Why can't the people who produced the technology, which has already shown that a different world is possible, just do the things that would prevent extinction?

SOCIAL Our A.I. Problem

Why can't the people who produced the technology, which has already shown that a different world is possible, just do the things that would prevent extinction?

SEO Our A.I. Problem

The fears expressed by A.I. executives are, mostly, genuine, Gideon Lewis-Kraus writes. Yet it's easy to see why they suffer from a credibility deficit. Surely it's the job of those creating the dangers to save us from them.