MIT Technology Review

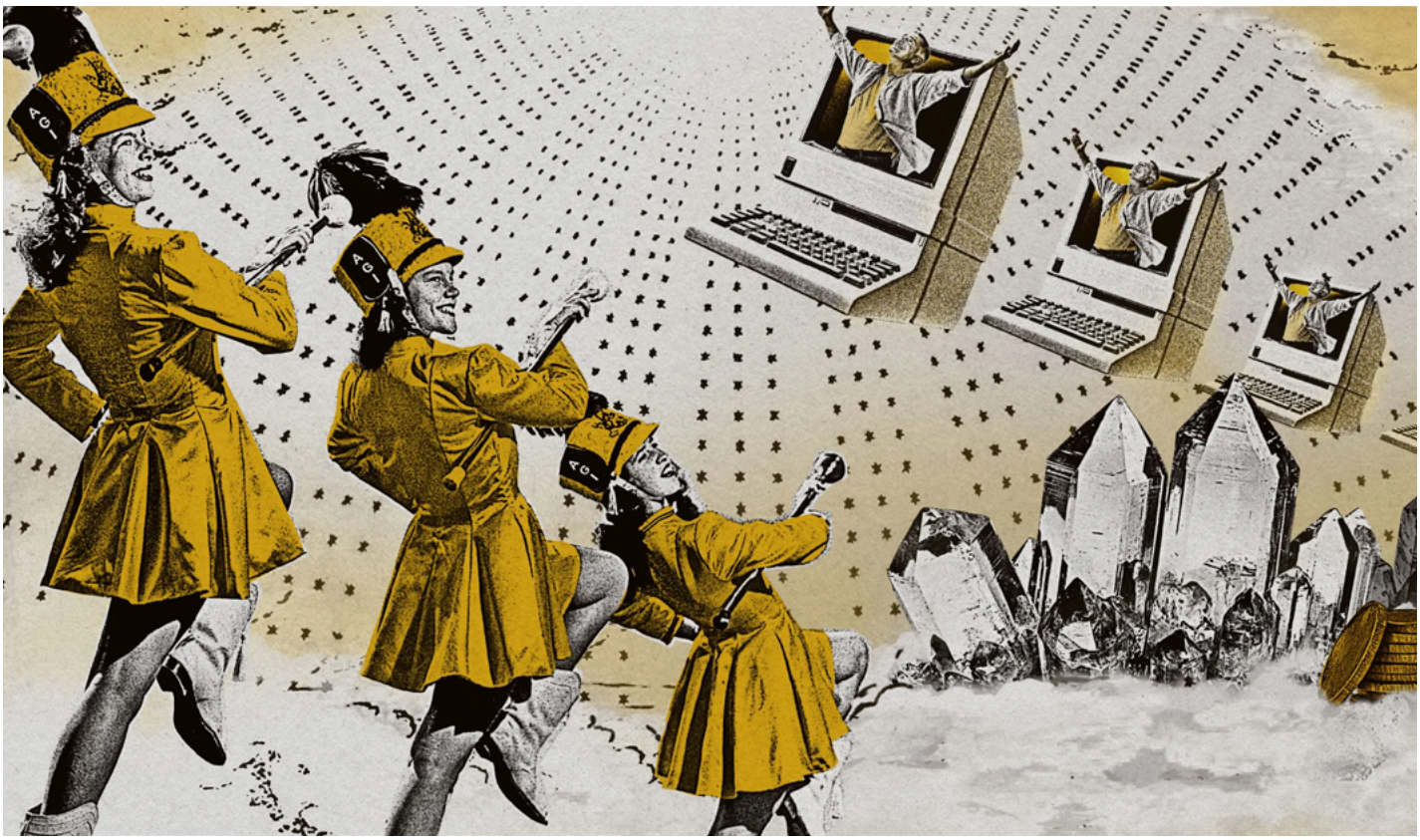
Subscriber Exclusive

ARTIFICIAL INTELLIGENCE

How AGI became the most consequential conspiracy theory of our time

The idea that machines will be as smart as – or smarter than – humans has hijacked an entire industry. But look closely and you'll see it's a myth that persists for many of the same reasons conspiracies do.

By **Will Douglas Heaven** October 30, 2025



STEPHANIE ARNETT/MIT TECHNOLOGY REVIEW | PUBLIC DOMAIN, ADOBE STOCK, GETTY IMAGES

Are you feeling it?

I hear it's close: two years, five years—maybe next year! And I hear it's going to change everything: it will cure disease, save the planet, and usher in an age of abundance. It will solve our biggest problems in ways we cannot yet imagine. It will redefine what it means to be human.

Wait—what if that's all too good to be true? Because I also hear it will bring on the apocalypse and kill us all...

Either way, and whatever your timeline, something big is about to happen.

We could be talking about the Second Coming. Or the day when Heaven's Gaters imagined they'd be picked up by a UFO and transformed into enlightened aliens. Or the moment when Donald Trump finally decides to deliver the storm that Q promised. But no. We're of course talking about artificial general intelligence, or AGI—that hypothetical near-future technology that (I hear) will be able to do pretty much whatever a human brain can do.

This story is part of MIT Technology Review*'s series "The New Conspiracy Age," on how the present boom in conspiracy theories is reshaping science and technology.*

For many, AGI is more than just a technology. In tech hubs like Silicon Valley, it's talked about in mystical terms. Ilya Sutskever, cofounder and former chief scientist at OpenAI, is said to have led chants of "Feel the AGI!" at team meetings. And he feels it more than most: In 2024, he left OpenAI, whose stated mission is to ensure that AGI benefits all of humanity, to cofound Safe Superintelligence, a startup dedicated to figuring out how to avoid a so-called rogue AGI (or control it when it comes). Superintelligence is the hot new flavor—*AGI but better!*—introduced as talk of AGI becomes commonplace.

Sutskever also exemplifies the mixed-up motivations at play among many self-anointed AGI evangelists. He has spent his career building the foundations for a future technology that he now finds terrifying. "It's going to be monumental, earth-shattering—there will be a before and an after," he told me a few months before he quit OpenAI. When I asked him why he had redirected his efforts into reining that technology in, he said: "I'm doing it for my own self-interest. It's obviously important that any superintelligence anyone builds does not go rogue. Obviously."

He's far from alone in his grandiose, even apocalyptic, thinking. Every age has its believers, people with an unshakeable faith that something huge is about to happen—a before and an after that they are privileged (or doomed) to live through.

For us, that's the promised advent of AGI. People are used to hearing that this or that is the next big thing, says Shannon Vallor, who studies the ethics of technology at the University of Edinburgh. "It used to be the computer age and then it was the internet age and now it's the AI age," she says. "It's normal to have something presented to you and be told that this thing is the future. What's different, of course, is that in contrast to computers and the internet, AGI doesn't exist."

And that's why *feeling the AGI* is not the same as boosting the next big thing. There's something weirder going on. Here's what I think: AGI is a lot like a conspiracy theory, and it may be the most consequential one of our time.

I have been reporting on artificial intelligence for more than a decade, and I've watched the idea of AGI bubble up from the backwaters to become the dominant narrative shaping an entire industry. A onetime pipe dream now props up the profit lines of some of the world's most valuable companies and thus, you could argue, the US stock market. It justifies dizzying down payments on the new power plants and data centers that we're told are needed to make the dream come true. Fixated on this hypothetical technology, AI firms are selling us hard.

Just listen to what the heads of some of those companies are telling us. AGI will be as smart as an entire "country of geniuses" (Dario Amodei, CEO of Anthropic); it will kick-start "an era of maximum human flourishing, where we travel to the stars and colonize the galaxy" (Demis Hassabis, CEO of Google DeepMind); it will "massively increase abundance and prosperity," even encourage people to enjoy life more and have more children (Sam Altman, CEO of OpenAI). That's some product.

Or not. Don't forget the flip side, of course. When those people are not shilling for utopia, they're saving us from hell. In 2023, Amodei, Hassabis, and Altman all put their names to a 22-word statement that read: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war." Elon Musk says AI has a 20% chance of annihilating humans.

"I've noticed recently that superintelligence, which I thought was a concept you definitely shouldn't mention if

© 2026 MIT Technology Review

you want to be taken seriously in public, is being thrown around by tech CEOs who are apparently planning to build it," says Katja Grace, lead researcher at AI Impacts, an organization that surveys AI researchers about their field. "I think it's easy to feel like this is fine. They also say it's going to kill us, but they're laughing while they say it."

You have to admit it all sounds a bit tinfoil hat. If you're building a conspiracy theory, you need a few things in the mix: a scheme that's flexible enough to sustain belief even when things don't work out as planned; the promise of a better future that can be realized only if believers uncover hidden truths; and a hope for salvation from the horrors of this world.

AGI just about checks all those boxes. The more you poke at the idea, the more it starts to look like a conspiracy. It's not, of course—not exactly. And I'm not drawing this parallel to dismiss the very real, often jaw-dropping results achieved by many people in this field, including (or especially) the AGI believers.

But by zooming in on things that AGI has in common with genuine conspiracies, I think we can bring the whole concept into better focus and reveal it for what it is: a techno-utopian (or techno-dystopian—pick your pill) fever dream that got its hooks into some pretty deep-seated beliefs that have made it hard to shake.

This isn't just a provocative thought experiment. It's important to question what we're told about AGI because buying into the idea isn't harmless. Right now, AGI is the most important narrative in tech—and, to some extent, in the global economy. We can't make sense of what's going on in AI without understanding where the idea of AGI came from, why it is so compelling, and how it shapes the way we think about technology overall.

I get it, I get it—calling AGI a conspiracy isn't a perfect analogy. It will also piss a lot of people off. But come with me down this rabbit hole and let me show you the light.

How Silicon Valley got AGI-pilled

It had a ring to it

A typical conspiracy theory usually starts out on the fringes. Maybe it's just a couple of people posting on a message board, gathering "evidence." Maybe it's a few people out in the desert with binoculars waiting to spot some bright lights in the sky. But some conspiracy theories get lucky, if you will: They start to percolate more widely; they start to become a bit more acceptable; they start to influence people in power. Maybe it's the UFOs (ahem, sorry, "unidentified aerial phenomena") that are now formally and openly discussed in government hearings. Maybe it's vaccine skepticism (yes, a much more dangerous example) that becomes official policy. And it's impossible to ignore that artificial general intelligence has followed a pretty similar trajectory to its more overtly conspiratorial brethren.

Let's go back to 2007, when AI wasn't sexy and it wasn't cool. Companies like Amazon and Netflix (which was still sending out DVDs in the mail) were using machine-learning models, proto-organisms to today's LLM behemoths, to recommend movies and books to customers. But that was more or less it.

Ben Goertzel had far bigger plans. About a decade earlier, the AI researcher had set up a dot-com startup called Webmind to train what he thought of as a kind of digital baby brain on the early internet. Childless, Webmind soon went bust.

But Goertzel was an influential figure in a fringe community of researchers who had dreamed for years of building humanlike artificial intelligence, an all-purpose computer program that could do many of the things people can do (and do them better). It was a vision that went far beyond the kind of tech that Netflix was experimenting with.

Goertzel wanted to put out a book promoting that vision, and he needed a name that would set it apart from the humdrum AI of the time. A former Webmind employee named Shane Legg suggested *Artificial General Intelligence*. It had a ring to it.

A few years later, Legg cofounded DeepMind with Demis Hassabis and Mustafa Suleyman. But to most serious researchers at the time, the claim that AI would one day mimic human abilities was a bit of a joke. AGI used to be a dirty word, Sutskever told me. Andrew Ng, founder of Google Brain and former chief scientist at the Chinese tech giant Baidu, told me he thought it was loony.

So what happened? I caught up with Goertzel last month to ask how a fringe idea went from crackpot to commonplace. "I'm sort of a complex chaotic systems guy, so I have a low estimate that I actually know what the nonlinear dynamic in the memosphere really was," he said. (Translation: It's complicated.)

Goertzel reckons a few things took the idea mainstream. The first is the Conference on Artificial General Intelligence, an annual meeting of researchers that

he helped set up in 2008, the year after his book was published. The conference was often coordinated with top mainstream academic meetups, such as the Association for the Advancement of Artificial Intelligence conference and the International Joint Conference on Artificial Intelligence. "If I just published a book with that name *AGI*, it possibly would have just come and gone," says Goertzel. "But the conference was circling through every year, with more and more students coming."

Next is Legg, who took the term with him to DeepMind. "I think they were the first mainstream corporate entity to talk about AGI," says Goertzel. "It wasn't the main thing they were harping on, but Shane and Demis would talk about it now and then. That was certainly a source of legitimation."

When I first talked to Legg about AGI five years ago, he said: "Talking about AGI in the early 2000s put you on the lunatic fringe … Even when we started DeepMind in 2010, we got an astonishing amount of eye-rolling at conferences." But by 2020 the wind had changed. "Some people are uncomfortable with it, but it's coming in from the cold," he told me.

The third thing Goertzel points to is the overlap between early AGI evangelists and Big Tech power brokers. In the years between shutting down Webmind and publishing that AGI book, Goertzel did some work with Peter Thiel at Thiel's hedge fund Clarium Capital. "We talked a bunch," says Goertzel. He recalls spending a day with Thiel at the Four Seasons in San Francisco. "I was trying to drum AGI into his head," says Goertzel. "But then he was also hearing from Eliezer how AGI is going to kill everybody."

Enter the doomers

That's Eliezer Yudkowsky, another influential figure who has done at least as much as Goertzel, if not more, to push the idea of AGI. But unlike Goertzel, Yudkowsky thinks there's a very high chance—99.5% is one number he throws out—that the development of AGI will be a catastrophe.

In 2000, Yudkowsky cofounded a nonprofit research outfit called the Singularity Institute for Artificial Intelligence (later renamed the Machine Intelligence Research Institute), which pretty quickly dedicated itself to preventing doomer scenarios. Thiel was an early benefactor.

At first, Yudkowsky's ideas didn't get much pickup. Recall that back then the idea of an all-powerful AI—let alone a dangerous one—was pure sci-fi. But in 2014, Nick Bostrom, a philosopher at the University of Oxford, published a book called *Superintelligence*.

"It put the AGI thing out there," says Goertzel. "I mean, Bill Gates, Elon Musk—lots of tech-industry AI people—read that book, and whether or not they agreed with his doomer perspective, Nick took Eliezer's concepts and wrapped them up in a very acceptable way."

"All of these things gave AGI a stamp of acceptability," Goertzel adds. "Rather than it being pure crackpot stuff from mavericks howling out in the wilderness."



STEPHANIE ARNETT/MIT TECHNOLOGY REVIEW | PUBLIC DOMAIN

Yudkowsky has been banging the same drum for 25 years; many engineers at today's top AI companies grew up reading and discussing his views online, especially on LessWrong, a popular hub for the tech industry's fervent community of rationalists and effective altruists.

Today, those views are more popular than ever, capturing the imagination of a younger generation of doomers like David Krueger, a researcher at the University of Montreal who previously served as research director at the UK's AI Security Institute. "I think we are definitely on track

to build superhuman AI systems that will kill everybody," Krueger tells me. "And I think that's horrible and we should stop immediately."

Yudkowsky gets profiled by the likes of the *New York Times*, which bills him as "Silicon Valley's version of a doomsday preacher." His new book, *If Anyone Builds It, Everyone Dies*, written with Nate Soares, president of the Machine Intelligence Research Institute, lays out wild claims, with little evidence, that unless we pull the plug on development, near-future AGI will lead to global Armageddon. The pair's position is extreme: They argue that an international ban should be enforced at all costs, up to and including the point of nuclear retaliation. After all, "datacenters can kill more people than nuclear weapons," Yudkowsky and Soares write.

This stuff is no longer niche. The book is an *NYT* bestseller and comes with endorsements from national security experts such as Suzanne Spaulding, a former US Department of Homeland Security official, and Fiona Hill, former senior director of the White House National Security Council, who now advises the UK government; celebrity scientists such as Max Tegmark and George Church; and other household names, including Stephen Fry, Mark Ruffalo, and Grimes. Yudkowsky now has a megaphone.

Still, it is those early quiet words in certain ears that may prove most consequential. Yudkowsky is credited with introducing Thiel to DeepMind's founders, after which Thiel became one of the first big investors in the company. Having merged with Google, it is now the in-house AI lab for the tech colossus Alphabet.

Alongside Musk, Thiel was also instrumental in setting up OpenAI in 2015, sinking millions into a startup founded on the singular ambition to build AGI—and make it safe. In 2023, OpenAI CEO Sam Altman posted on X: "eliezer has IMO done more to accelerate AGI than anyone else. certainly he got many of us interested in AGI." Yudkowsky might one day deserve the Nobel Peace Prize for that, Altman added. But by this point, Thiel had apparently grown wary of the "AI safety people" and the power they were gaining. "You don't understand how Eliezer has programmed half the people in your company to believe in that stuff," he is reported to have told Altman at a dinner party in late 2023. "You need to take this more seriously." Altman "tried not to roll his eyes," according to *Wall Street Journal* reporter Keach Hagey.

OpenAI is now the most valuable private company in the world, worth half a trillion dollars.

And the transformation is complete: Like all the most powerful conspiracies, AGI has slipped into the mainstream and taken hold.

The great AGI conspiracy

The term "AGI" may have been popularized less than 20 years ago, but the mythmaking behind it has been there since the start of the computer age—a cosmic microwave background of chutzpah and marketing.

Alan Turing asked if machines could think only five years after the first electronic computer, ENIAC, was built in 1945. And here's Turing a little later, in a 1951 radio broadcast: "It seems probable that once the machine thinking method had started, it would not take long to outstrip our feeble powers. There would be no question of the machines dying, and they would be able to converse with each other to sharpen their wits. At some stage therefore we should have to expect the machines to take control."

Then, in 1955, the computer scientist John McCarthy and his colleagues applied for US government funding to create what they fatefully chose to call "artificial intelligence"—a canny spin, given that computers at the time were the size of a room and as dumb as a thermostat. Even so, as McCarthy wrote in that funding application: "An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves."

It's this myth that's the root of the AGI conspiracy. A smarter-than-human machine that can do it all is not a technology. It's a dream, unmoored from reality. Once you see that, other parallels with conspiracy thinking start to leap out.

It's impossible to debunk a shape-shifting idea like AGI.

Talking about AGI can sometimes feel like arguing with an enthusiastic Redditor about what drugs (or particles in the sky) are controlling your mind. Each point has a counterpoint that tries to chip away at your own sense of what's true. Ultimately, it's a clash of worldviews, not an exchange of evidence-based reason. AGI is like that, too—it's slippery.

Part of the issue is that despite all the money, all the talk, nobody knows how to build it. More than that: Most people don't even agree on what AGI really is—which

helps explain how people can get away with telling us it can both save the world and end it. At the core of most definitions you'll find the idea of a machine that can match humans on a wide range of cognitive tasks. (And remember, superintelligence is AGI's shiny new upgrade: a machine that can *outmatch* us.) But even that's easy to pull apart: What humans are we talking about? What kind of cognitive task? And how wide a range?

"There's no real definition of it," says Christopher Symons, chief artificial intelligence scientist at the AI health-care startup Lirio and former head of the computer science and math division at Oak Ridge National Laboratory. "If you say 'human-level intelligence,' that could be an infinite number of things—everybody's level of intelligence is slightly different."

And so, says Symons, we're in this weird race to build … what, exactly? "What are you trying to get it to do?"

In 2023, a team of researchers at Google DeepMind, including Legg, had a go at categorizing various definitions that people had proposed for AGI. Some said that a machine had to be able to learn; some said that it had to be able to make money; some said that it had to have a body and move about in the world (and maybe make coffee).

Legg told me that when he'd suggested the term to Goertzel for the title of his book, the hand-waviness had been kind of the point. "I didn't have an especially clear definition. I didn't really feel it was necessary," he said at the time. "I was actually thinking of it more as a field of study, rather than an artifact."

So, I guess we'll know it when we see it? The problem is that some people think they've seen it already.

In 2023, a team of Microsoft researchers put out a paper in which they described their experiences playing around with a prerelease version of OpenAI's large language model GPT-4. They called it "Sparks of Artificial General Intelligence"—and it polarized the industry.

It was a moment when a lot of researchers were blown away and trying to come to terms with what they were seeing. "Shit was working better than they had expected it to," says Goertzel. "The concept of AGI genuinely started to seem more plausible."

And yet for all of LLMs' remarkable wordplay, Goertzel doesn't think that they do in fact contain sparks of AGI. "It's a little surprising to me that some people with a deep technical understanding of how these tools work under the hood still think that they could become human-level AGI," he says. "On the other hand, you can't prove it's not true."

And there it is: *You can't prove it's not true*. "The idea that AGI is coming and that it's right around the corner and that it's inevitable has licensed a great many departures from reality," says the University of Edinburgh's Vallor. "But we really don't have any evidence for it."

Conspiracy thinking looms again. Predictions about when AGI will arrive are made with the precision of numerologists counting down to the end of days. With no real stakes in the game, deadlines come and go with a shrug. Excuses are made and timelines are adjusted yet again.

We saw this when OpenAI released the much-hyped GPT-5 this summer. AI stans were disappointed that the new version of the company's flagship technology wasn't the step change they expected. But instead of seeing that as evidence that AGI wasn't attainable—or attainable with an LLM, at least—believers pushed out their predictions for how soon AGI would come. It was coming—just, you know, next time.

Maybe they're right. Or maybe people will pick whatever evidence they can to defend an idea and overlook evidence that counts against it. Jeremy Cohen, who studies conspiracy thinking in technology circles at McMaster University in Canada, calls this imperfect evidence gathering—a hallmark of conspiracy thinking.

Cohen started his research career in the Arizona desert, studying a community called People Unlimited that believed its members were immortal. The conviction was impervious to contrary evidence. When its members died of natural causes (including two of its founders), the thinking was that they must have deserved it. "The general consensus was that every death was a suicide," says Cohen. "If you are immortal and you get cancer and you die—well, you must have done something wrong."

Cohen has since been focused on transhumanism (the idea that technology can help humans push past their natural limitations) and AGI. "I am seeing a lot of parallels. There are forms of magical thinking that I think is a part of the popular imagination around AGI," he says. "It connects really well to the kinds of religious imaginaries that you see in conspiracy thinking today."

The believers are in on the AGI secret.

Maybe some of you think I'm an idiot: *You don't get it at all lol*. But that's kind of my point. There are insiders and

outsiders. When I talk to researchers or engineers who are happy to drop AGI into the conversation as a given, it's like they know something I don't. But nobody's ever been able to tell me what that something is.

The truth is out there, if you know where to look. Conspiracy theories are primarily concerned about revealing a hidden truth, Cohen tells me: "It's a really fundamental part of conspiracy thinking, and that's absolutely something that you see in the way people talk about AGI," he says.

Last year, a 23-year-old former OpenAI staffer turned investor, Leopold Aschenbrenner, published a much-dissected 165-page manifesto titled "Situational Awareness." You don't need to read it to get the idea: You either see the truth of what's coming or you don't. And you don't need cold, hard facts, either—it's enough to *feel* it. Those who don't just haven't seen the light.

This idea stalked the periphery of my conversation with Goertzel, too. When I pushed him on why people are skeptical of AGI, for instance, he said: "Before every major technical achievement, from human flight to electrical power, loads of wise pundits would tell you why it was never going to happen. The fact is, most people only believe what they see in front of their faces."

That makes AGI sound like an article of faith. I put that to Krueger, who believes AGI's arrival is maybe five years out. He scoffed: "I think that's completely backwards." For him, the article of faith is the idea that it *won't* happen—it's the skeptics who continue to deny the obvious. (Even so, he hedges: No one knows for sure, he says, but there's no obvious reason that AGI won't come.)

Hidden truths bring truth seekers, bent on revealing what they've been able to see all along. With AGI, though, it's not enough to *uncover* something hidden. Here, revelation requires an unprecedented act of *creation*. If you believe AGI is achievable, then you believe that those making it are midwives to machines that will match or surpass human intelligence. "The idea of giving birth to machine gods is obviously very flattering to the ego," says Vallor. "It's an incredibly seductive thing to think that you yourself are laying the early foundations for that transcendence."

It's yet another overlap with conspiracy thinking. Part of the draw is the desire for a sense of purpose in an otherwise messy world that can feel meaningless—the longing to be a person of consequence.

Krueger, who is based in Berkeley, says he knows people working on AI who see the technology as our natural successor. "They view it as akin to having children or something," he says. "Side note: they usually don't have children."

AGI will be our one true savior (or it'll bring the apocalypse).

Cohen sees parallels between many modern conspiracy theories and the New Age movement, which reached its peak of influence in the 1970s and '80s. Adherents believed humanity was on the cusp of unlocking an era of spiritual well-being and expanded consciousness that would usher in a more peaceful and prosperous world. In a nutshell, the idea was that by engaging in a set of pseudo-religious practices, including astrology and the careful curation of crystals, humans would transcend their limitations and enter a kind of hippie utopia.

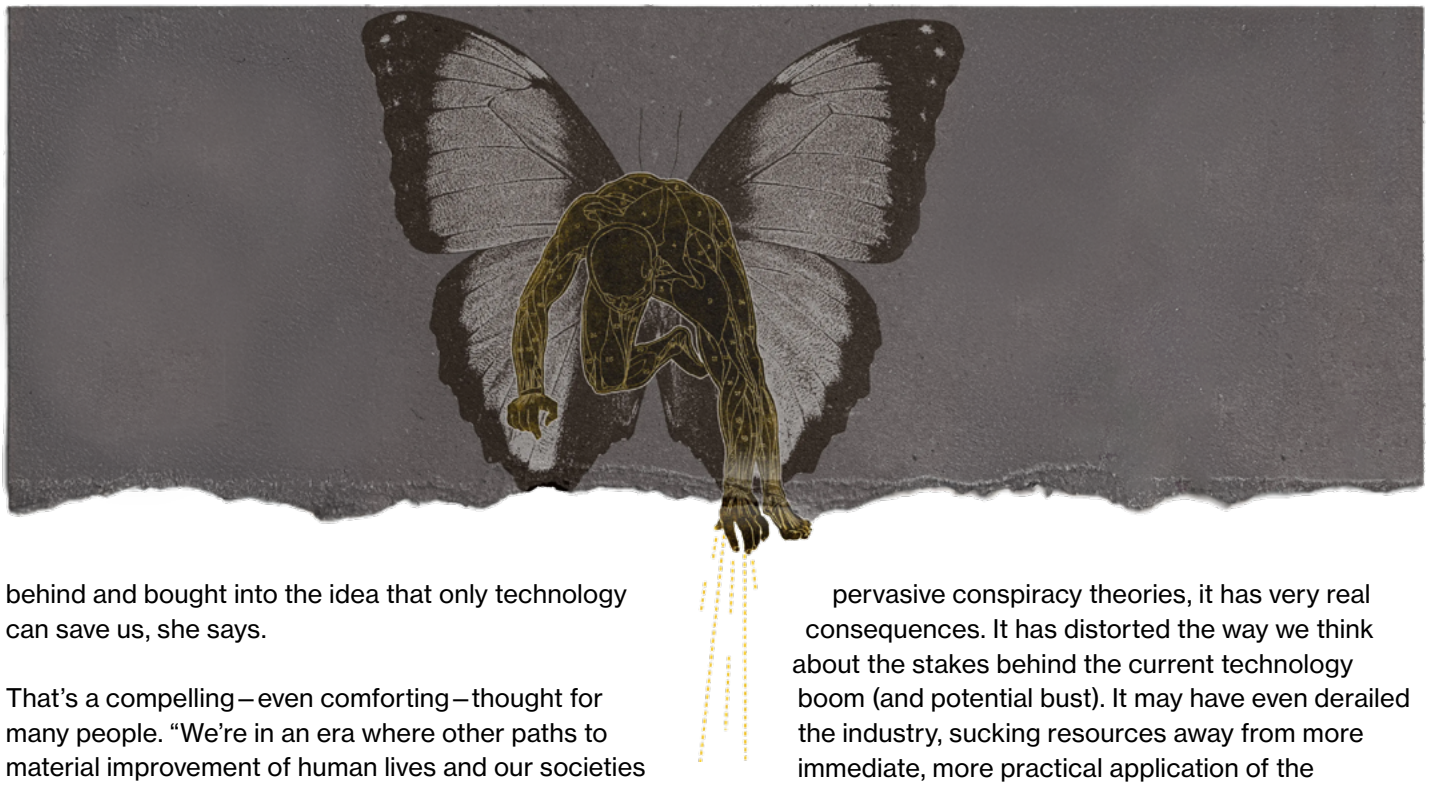
Today's tech industry is built on compute, not crystals, but its sense of what's at stake is no less transcendent: "You know, this idea that there is going to be this fundamental shift, there's going to be this millenarian turn where we end up in a techno-utopian future," says Cohen. "And the idea that AGI is going to ultimately allow humanity to overcome the problems that face us."

In many people's telling, AGI will arrive all at once. Incremental advances in AI will stack up until, one day, AI will be good enough to start making better AI by itself. At which point—*FOOM*—it will advance so rapidly that AGI will arrive in what's often called an intelligence explosion, leading to a point of no return known as the Singularity, a goofy term that's been popular in AGI circles for years. Co-opting a concept from physics, the science fiction author Vernor Vinge first introduced the idea of a technological singularity in the 1980s. Vinge imagined an event horizon on the path of technological progress beyond which humans would be fast outstripped by the exponential self-improvement of the machines they had created.

Call it the AI Big Bang—which, again, gives us a before and an after, a transcendent moment when humanity as we know it changes forever (for good or bad). "People imagine it as an event," says Grace from AI Impacts.

For Vallor, this belief system is notable for the way that a faith in technology has replaced a faith in humans. Despite the woo-woo, New Age thinking was at least motivated by the idea that people had what it took to change the world by themselves, if they could only tap into it. With the pursuit of AGI, we've left that self-belief

© 2026 MIT Technology Review



behind and bought into the idea that only technology can save us, she says.

That's a compelling—even comforting—thought for many people. "We're in an era where other paths to material improvement of human lives and our societies seem to have been exhausted," Vallor says.

Technology once promised a route to a better future: Progress was a ladder that we would climb toward human and social flourishing. "We've passed the peak of that," says Vallor. "I think the one thing that gives many people hope and a return to that kind of optimism about the future is AGI."

Push this idea to its conclusion and, again, AGI becomes a kind of god—one that can offer relief from earthly suffering, says Vallor.

Kelly Joyce, a sociologist at the University of North Carolina who studies how cultural, political, and economic beliefs shape the way we think about and use technology, sees all these wild predictions about AGI as something more banal: part of a long-term pattern of overpromising from the tech industry. "What's interesting to me is that we get sucked in every time," she says. "There is a deep belief that technology is better than human beings."

Joyce thinks that's why, when the hype kicks in, people are predisposed to believe it. "It's a religion," she says. "We believe in technology. Technology is God. It's really hard to push back against it. People don't want to hear it."

How AGI hijacked an industry

The fantasy of computers that can do almost anything a person can is seductive. But like many pervasive conspiracy theories, it has very real consequences. It has distorted the way we think about the stakes behind the current technology boom (and potential bust). It may have even derailed the industry, sucking resources away from more immediate, more practical application of the technology. More than anything else, it gives us a free pass to be lazy. It fools us into thinking we might be able to avoid the actual hard work needed to solve intractable, world-spanning problems—problems that will require international cooperation and compromise and expensive aid. Why bother with that when we'll soon have machines to figure it all out for us?

Consider the resources being sunk into this grand project. Just last month, OpenAI and Nvidia announced an up-to-$100 billion partnership that would see the chip giant supply at least 10 gigawatts of ChatGPT's insatiable demand. That's higher than nuclear power plant numbers. A bolt of lightning might release that much energy. The flux capacitor inside Dr. Emmett Brown's DeLorean time machine only required 1.2 gigawatts to send Marty back to the future. And then, only two weeks later, OpenAI announced a second partnership with chipmaker AMD for another six gigawatts of power.

Promoting the Nvidia deal on CNBC, Altman, straight-faced, claimed that without this kind of data center buildout, people would have to choose between a cure for cancer and free education. "No one wants to make that choice," he said. (Just a few weeks later, he announced that erotic chats would be coming to ChatGPT.)

Add to those costs the loss of investment in more immediate technology that could change lives today and tomorrow and the next day. "To me it's a huge missed opportunity," says Lirio's Symons, "to put all these

resources into solving something nebulous when we already know there's real problems that we could solve."

But that's not how the likes of OpenAI needs to operate. "With people throwing so much money at these companies, they don't have to do that," Symons says. "If you've got hundreds of billions of dollars, you don't have to focus on a practical, solvable project."

Despite his steadfast belief that AGI is coming, Krueger also thinks the industry's single-minded pursuit of it means that potential solutions to real problems, such as better health care, are being ignored. "People have a long list of complaints about both the concept of AGI and the idea that it should be a goal," he says. "I think it's pretty unpopular in the field."

And there are consequences for the way governments support and regulate technology (or don't). Tina Law, who studies technology policy at the University of California–Davis, worries that policymakers are getting lobbied about the ways AI will one day kill us all, instead of addressing real concerns about the ways AI could impact people's lives in immediate and material ways today. Inequality has been sidetracked by existential risk.

"Hype is a lucrative strategy for tech firms," says Law. A big part of that hype is the idea that what's happening is inevitable: If we don't build it, someone else will. "When something is framed as inevitable," Law says, "people doubt not only whether they should resist but also whether they have the capacity to do so." Everyone gets locked in.

The AGI distortion field isn't limited to tech policy, says Milton Mueller at the Georgia Institute of Technology, who works on technology policy and regulation. The race to AGI gets compared to the race to the atomic bomb, he says. "So whoever gets it first is going to have ultimate power over everybody else. That's a crazy and dangerous idea that really will distort our approach to foreign policy."

There's a business incentive for companies (and governments) to push the myth of AGI, says Mueller, because they can then claim that they will be the first to get there. But because they're running a race in which nobody has agreed on the finish line, the myth can be spun as long as it's useful. Or as long as investors are willing to buy into it.

It's not hard to see how this plays out. It's not utopia or hell—it's OpenAI and its peers making a whole lot more money.

The great AGI conspiracy, concluded

And maybe that brings us back to the whole conspiracy thing—and a late-game twist in this tale. So far we've ignored one popular feature of conspiracy thinking: that there's a group of powerful figures pulling the levers behind the scenes and that, by seeking the truth, believers can expose this elite cabal.

Sure, the people feeling the AGI aren't publicly accusing any Illuminati or WEF-like force of preventing the AGI future or withholding its secrets.

But what if there are, in fact, shadowy puppet masters here—and they're the very people who have pushed the AGI conspiracy hardest all along? The kings of Silicon Valley are throwing everything they can get at building AGI for profit. The myth of AGI serves their interests more than anybody else's.

As one senior executive at an AI company said to us recently, AGI always needs to be six months to a year away, because if it's any further than that, you won't be able to recruit people from Jane Street, and if it's closer to already here, then what's the point?

As Vallor puts it: "If OpenAI says they're building a machine that's going to make corporations even more powerful than they are today, that isn't going to get the kind of public buy-in that they need."

Remember: You create a god and you become like one yourself. Krueger says there's a line of thinking running through Silicon Valley in which building AI is a way to seize huge amounts of power. (It's one of the premises of Aschenbrenner's "Situational Awareness," for example.) "You know, we're going to have this godlike power and we're going to have to figure out what to do with it," says Krueger. "A lot of people think if they get there first, they can basically take over the world."

"They're putting so much effort into selling their vision of a future with AGI in it, and they're having a pretty good amount of success because they have so much power," he adds.

Goertzel, for one, is almost lamenting how successful the maybe-cabal has been. He's actually starting to miss life on the fringes. "In my generation, you had to have a lot of vision to want to work on AGI, and you had to be very stubborn," he says. "Now it's almost, like, what your

grandma tells you to do to get a job instead of being a business major."

"It's disorienting that this stuff is so broadly accepted," he says. "It almost gives me the desire to go work on something else that not so many people are doing." He's half joking (I think): "Obviously, putting the finishing touches to AGI is more important than gratifying my preference to be out on the frontier."

But I'm no clearer on what exactly they're putting the finishing touches on. What does it mean for technology in general if we fall so hard for the fairy tales? In a lot of ways, I think the whole idea of AGI is built on a warped view of what we should expect technology to do, and even what intelligence is in the first place. Stripped back to its essentials, the argument for AGI rests on the premise that one technology, AI, has gotten very good, very fast, and will continue to get better. But set aside the technical objections—what if it doesn't continue to get better?—and you're left with the claim that intelligence is a commodity you can get more of if you have the right data or compute or neural network. And it's not.

Intelligence doesn't come as a quantity you can just ratchet up and up. Smart people may be brilliant in one area and not in others. Some Nobel Prize winners are really bad at playing the piano or caring for their kids. Some very smart people insist that AGI is coming next year.

It's hard not to wonder what will get its hooks into us next.

Before we ended our call, Goertzel told me about an event he'd just been to in San Francisco on AI consciousness and parapsychology: "ESP, precognition, and whatnot."

"That's where AGI was 20 years ago," he said. "Everyone thinks it's batshit crazy."

Correction: This story has been updated to better reflect David Krueger's views.

Copying or posting this work is an infringement of copyright.