

Jun 3, 2025 - Technology

AI pioneer Bengio launches \$30M nonprofit to rethink safety



Ina Fried



Yoshua Bengio, testifying before a Senate subcommittee in 2023. Photo: Alex Wong/Getty Images

Machine learning pioneer Yoshua Bengio is launching a new nonprofit lab backed by roughly \$30 million in funding to make [AI](#) systems act less like humans.

Why it matters: The move bucks a trend toward AI that acts independently, which Bengio and others fear might create systems that place their own interests over humanity's.

- "We've been getting inspiration from humans as the template for building intelligent machines, but that's crazy, right?" Bengio said in an interview.
- "If we continue on this path, that means we're going to be creating entities — like us — that don't want to die, and that may be smarter than us and that we're not sure if they're going to behave according to our norms and our instructions," he said.

Driving the news: Bengio, a Montreal-based researcher who has [long warned about the risks](#) of a technology he helped develop, has raised about \$30 million for the nonprofit, dubbed LawZero.

- LawZero currently has about 15 staffers, Bengio said, "but the goal is to hire many more."
- Bengio is among those who have [called for tougher regulation](#) of AI development and even [the breakup of big tech companies](#).
- Earlier this year he [gave a TED Talk](#) urging greater caution and collaboration.

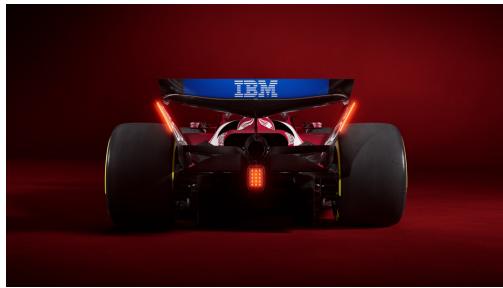
The big picture: There's a growing sense of worry among critics — and even AI practitioners — that safety is taking a back seat as companies and countries race to be first with AI that can best humans in a wide variety of tasks, so-called artificial general intelligence (AGI).

- Bengio said there is also a high risk in concentrating control of advanced AI in a handful of companies.
- "You don't want AGI or superintelligence to be in the hands of one person or one company only deciding what to do, or even one government," Bengio said. "So

you need very strong checks and balances."

Between the lines: Bengio says a large part of the problem is how current systems are trained. During initial training, the systems are taught to mimic humans, and then they're honed by seeing which responses people find most appealing.

- "Both of these give rise to uncontrolled agency," Bengio said.



A message from IBM

Scuderia Ferrari rebuilt its app with IBM Watsonx to deliver an AI-powered fan experience.

[Learn more →](#)

- Some early glimmers of this are already appearing — as when Anthropic's latest model, in a test scenario, sought to blackmail its engineers to avoid shutdown.
- By contrast, Bengio says he wants to create AI systems that have intellectual distance from humans and act as more of a detached scientist than a personal companion or human agent.
- "The training principle is completely different, but it can exploit a lot of the recent advances that have happened in machine learning," he said.

Yes, but: Bengio told Axios that the \$30 million should be enough to fund the basic

research effort for about 18 months.

- But AI is expensive, and Bengio's new nonprofit may find it difficult to raise additional, larger funding rounds — as OpenAI and others have found.
- Bengio says he doesn't see future funding as a roadblock, as he's certain more investors now understand the plausible risks ahead.
- Governments could also be future backers of LawZero, Bengio said.



Want more stories like this? Sign up for [Axios Future of Defense](#)

Enter your email address

Subscribe

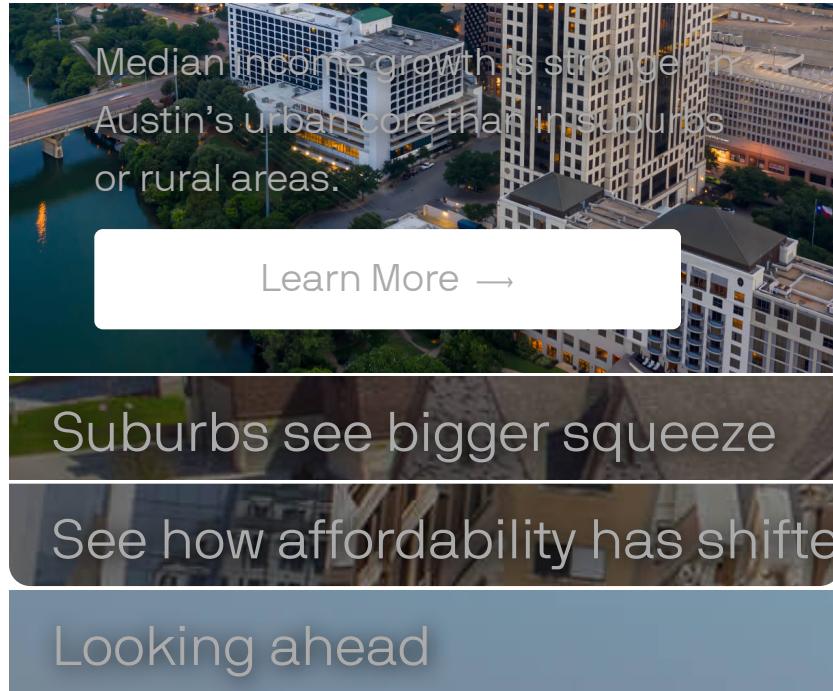
Advertisement

A message from **JPMorganChase**

An advertisement banner for JPMorgan Chase. It features a large image of a city skyline at dusk or night. Overlaid on the image are three text snippets in different colors:

- The affordability gap widens (white text)
- Employment more than doubles (light blue text)
- City center holds its ground (yellow/gold text)

The overall theme suggests economic trends and stability in the city.



Go deeper



Barak Ravid
2 hours ago - World

State Department considers giving \$500 million to the new Gaza Humanitarian Foundation

Palestinian children wait as charitable organizations distribute hot meals in the Nuseirat refugee camp, on June 2 amid a severe humanitarian crisis in the region. Photo: Moiz Salhi/Anadolu via Getty Images

The State Department is discussing allocating \$500 million for funding of the U.S. and Israel backed [Gaza Humanitarian Foundation \(GHF\)](#) that has been delivering food to Palestinians in Gaza in recent weeks, a U.S. official and a source with direct knowledge told Axios.

Why it matters: If the U.S. goes ahead with this massive funding, it will become the biggest donor to the foundation and will de facto "own" the operation.

[Go deeper \(2 min. read\) →](#)



Sara Fischer
3 hours ago - Business

Appeals court allows White House AP ban to continue

Photo: Artur Widak/NurPhoto via Getty Images

A panel of judges from a U.S. federal appeals court on Friday said parts of the White House's ban on the [Associated Press](#) could remain, dealing a devastating blow to the AP.

Why it matters: Press freedom advocates are closely watching the AP's case for any precedents it could set around free speech protections for journalists.

[Go deeper \(1 min. read\) →](#)



Hans Nichols
4 hours ago - Politics & Policy

Scoop: Treasury leads negotiation as "revenge tax" looms

Treasury Secretary Scott Bessent speaks during a meeting with President Trump and German Chancellor Friedrich Merz.
Photo: Chris Kleponis/CNP/Bloomberg via Getty Images

Top Treasury officials are privately explaining to GOP senators [that Section 899](#) of the House-passed budget bill is already forcing foreign countries to the negotiating table, according to administration officials.

Why it matters: Critics are calling the [provision a "revenge tax."](#) But the Trump administration sees Section 899 as an important tool — like tariffs — to help negotiate better deals for American multinational corporations.

[Go deeper \(2 min. read\) →](#)



Smarter, faster on what matters.

[Explore Axios Newsletters](#)

[About Axios](#)

[Axios Live](#)

[Advertise with us](#)

[Axios HQ](#)

[Careers](#)

[Privacy policy](#)

[Contact us](#)

[Terms of use](#)

[Newsletters](#)

[Your Privacy Choices](#)

[Axios Entertainment](#)

Axios Media Inc., 2025