

[technologyreview.com](https://www.technologyreview.com)

One of the fathers of AI is worried about its future

Will Knight

5-7 minutes

Yoshua Bengio is a grand master of modern artificial intelligence.

Alongside [Geoff Hinton](#) and [Yann LeCun](#), Bengio is famous for championing a technique known as [deep learning](#) that in recent years has gone from an academic curiosity to one of the most powerful technologies on the planet.

Deep learning involves feeding data to large neural networks that crudely simulate the human brain, and it has proved incredibly powerful and effective for all sorts of practical tasks, from voice recognition and image classification to controlling self-driving cars and automating business decisions.

Bengio has resisted the lure of any big tech company. While Hinton and LeCun joined Google and Facebook, respectively, he remains a full-time professor at the University of Montreal. (He did, however, cofound Element AI in 2016, and it has built a very successful business helping big companies explore the commercial applications of AI research.)

Bengio met with MIT Technology Review's senior editor for AI, Will Knight, at an MIT event recently.

What do you make of the idea that there's an AI race between different countries?

I don't like it. I don't think it's the right way to do it.

We could collectively participate in a race, but as a scientist and somebody who wants to think about the common good, I think we're better off thinking about how to both build smarter machines and make sure AI is used for the well-being of as many people as possible.

Are there ways to foster more collaboration between countries?

We could make it easier for people from developing countries to come here. It is a big problem right now. In Europe or the US or Canada it is very difficult for an African researcher to get a visa. It's a lottery, and very often they will use any excuse to refuse access. This is totally unfair. It is already hard for them to do research with little resources, but in addition if they can't have access to the community, I think that's really unfair. As a way to counter some of that, we are going to have the ICLR conference [a major AI conference] in 2020 in Africa.

Inclusivity has to be more than a word we say to look good. The potential for AI to be useful in the developing world is even greater. They need to improve technology even more than we do, and they have different needs.

Are you worried about just a few AI companies, in the West and perhaps China, dominating the field of AI?

Yes, it's another reason why we need to have more democracy in AI research. It's that AI research by itself will tend to lead to

concentrations of power, money, and researchers. The best students want to go to the best companies. They have much more money, they have much more data. And this is not healthy. Even in a democracy, it's dangerous to have too much power concentrated in a few hands.

**There has been a lot of controversy over military uses of AI.
Where do you stand on that?**

I stand very firmly against.

Even non-lethal uses of AI?

Well, I don't want to prevent that. I think we need to make it immoral to have killer robots. We need to change the culture, and that includes changing laws and treaties. That can go a long way.

Of course, you'll never completely prevent it, and people say, "Some rogue country will develop these things." My answer is that one, we want to make them feel guilty for doing it, and two, there's nothing to stop us from building defensive technology. There's a big difference between defensive weapons that will kill off drones, and offensive weapons that are targeting humans. Both can use AI.

Shouldn't AI experts work with the military to ensure this happens?

If they had the right moral values, fine. But I don't completely trust military organizations, because they tend to put duty before morality. I wish it was different.

What are you most excited about in terms of new AI research?

I think we need to consider the hard challenges of AI and not be satisfied with short-term, incremental advances. I'm not saying I want to forget deep learning. On the contrary, I want to build on it.

But we need to be able to extend it to do things like reasoning, learning causality, and exploring the world in order to learn and acquire information.

If we really want to approach human-level AI, it's another ball game. We need long-term investments, and I think academia is the best place to carry that torch.

You mention causality—in other words, grasping not just patterns in data but *why* something happens. Why is that important, and why is it so hard?

If you have a good causal model of the world you are dealing with, you can generalize even in unfamiliar situations. That's crucial. We humans are able to project ourselves into situations that are very different from our day-to-day experience. Machines are not, because they don't have these causal models.

We can hand-craft them, but that's not enough. We need machines that can discover causal models. To some extent it's never going to be perfect. We don't have a perfect causal model of the reality; that's why we make a lot of mistakes. But we are much better off at doing this than other animals.

Right now, we don't really have good algorithms for this, but I think if enough people work at it and consider it important, we will make advances.

