

As Artificial Intelligence (AI) races ahead, Pope Leo's warning sparks a crucial debate: should technology replace human potential, or empower people to thrive alongside intelligent machines?

BY DARON ACEMOGLU

Artificial Intelligence (AI) is reshaping how we communicate, access information and work, how income and status are distributed and even how we wage war. Yet the public conversation remains narrowly focused on the competition between AI labs or abstract debates about the technology's capabilities. Almost no one is asking what purpose AI ought to serve, or whether our current mindset, institutions and control mechanisms are capable of steering the technology towards broad-based improvements in human welfare.

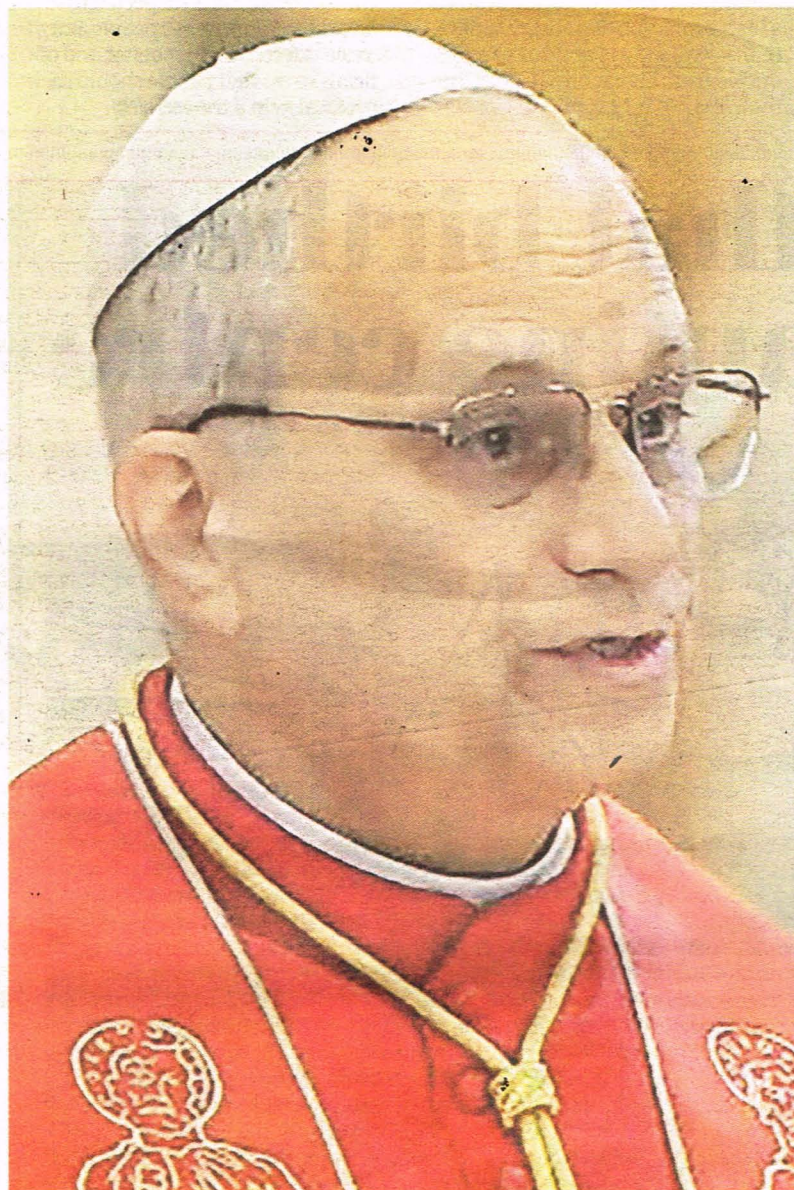
It was, therefore, refreshing to see Pope Leo XIV weigh in on the issue with his first encyclical, which describes AI's current trajectory as a profound threat to human dignity. As an economist who has long argued that technologically driven outcomes are matters of choice, not fate, I welcome his intervention.

Pope Leo is ahead of most commentators in pointing out that "technology is never neutral because it takes on the characteristics of those who devise, finance, regulate and use it". And yet, I worry that even he has not gone far enough on the most consequential question: What should AI be designed to do?

As Simon Johnson and I stress in our book *Power and Progress: Our Thousand-Year Struggle over Technology and Prosperity*, there are multiple paths that a technology such as AI can take, and each has far-reaching implications for society. For example, the Pope is right to question the current trajectory of AI in warfare and law enforcement. What was taboo only a few years ago -- AI-driven mass surveillance, algorithms selecting targets for killing -- has become routine.

With many in Silicon Valley urging the United States to reinforce its hard power through a new military-algorithmic complex, Pope Leo cautions that "Any technology that facilitates attacks without seeing the face of human beings lowers the moral thresh-

Pope should have gone further on AI – expert



Pope Leo XIV speaks at the Cathedral of Santa Maria Assunta in Acerra, Italy, on May 23. On Saturday, the Pope weighed in on AI's profound threat to human dignity. PHOTO/REUTERS

old of conflict". The Pope then calls for "disarmament of AI" to free "it from the mentality of 'armed' competition, which today is not limited simply to the military context, but is also an economic and cognitive phenomenon".

A more fundamental piece of wisdom underlies these specific concerns: technological progress is not necessarily moral progress. Just because something is technically feasible does not mean that it is good for human-

ity. Whether a technology is desirable depends on who controls it and on the ideology and interests that guide them.

Pope Leo does hint at what I see as the most immediate risk, namely that "while AI promises to boost productivity by taking over mundane tasks, it frequently forces workers to adapt to the speed and demands of machines, rather than designing machines to work with those who work".

AI design philosophy

But the Pope stops short of questioning the prevailing AI design philosophy. The entire AI industry's approach is centred on mimicking human capabilities and automating human tasks, with the goal of creating an "artificial general intelligence" that can do everything a person can.

This philosophy rests on the mistaken assumption that machine intelligence and human intelligence are fundamentally similar. Humans are "one-shot" learners. We form hypotheses from a few examples, simulate possibilities in our minds and refine our understanding through a social process of trial and error. Hence, children learn language by imitating a few words, generalising and adjusting their speech based on how others respond. We are not very good at absorbing massive volumes of information or sifting through unstructured data for relevant patterns.

By contrast, AI models thrive on enormous training sets and excel at pattern recognition at scale, but they have yet to demonstrate genuine creativity. They have no experience of real-world embodiment, nor any capacity for trial-and-error learning through interactions with the physical and social world (except in a limited way when there are clear rewards for reinforcement learning in specific domains).

When two things are different, you should not -- and typically cannot -- use one to mimic the other. The results would be suboptimal. It would have been a colossal mistake if Phil Jackson, the legendary coach of the Chicago Bulls in the 1990s, had pushed Michael Jordan to mimic everything that Scottie Pippen and Dennis Rodman were doing. The team went from championship to championship precisely because these players worked together and complemented each other's skills.

The same applies to AI and human skills. Using AI to do what humans cannot do, so that humans can expand what they do, is more productive than mimicry. In a future scenario where AI increases, rather than displaces, human capabilities, electricians would be aided by AI diagnostics, nurses would consult AI in interpreting

symptoms, and teachers might use AI to personalise instruction for each student.

Optimists and industry insiders might respond that automation-first AI can still benefit everyone, provided that redistributive policies keep pace. But this argument has a poor track record. Four decades of digital automation have already concentrated gains at the top, hollowed out middle-skill work and produced disappointing aggregate productivity growth. There is little reason to expect that an even more powerful round of automation, deployed by an even more concentrated industry, will end differently.

And the global stakes are even greater than those in the US. For billions of people in the developing world, where a decent job is the only reliable path out of poverty, an automation-centric AI agenda is a recipe for disaster. We can and must demand a different design.

Perhaps the biggest failing of today's AI industry is its refusal to recognise any of this. The handful of people unleashing this technology on the world are guided by an ideology of control (over humanity) and by a conviction that machines are uniformly better than humans.

Moral clarity

Pope Leo is right to call for moral clarity and a serious, society-wide debate. But the conversation must move beyond exhortation toward concrete choices: antitrust action against the dominant platforms, public investments in human-complementary AI, regulation of surveillance and autonomous weapons, and meaningful rights for workers and citizens over the data on which these systems depend.

Pope Leo's intervention makes such a response a little more likely than it was before. But the rest of us must stand up for humanity, too.

©2026 Project Syndicate

Daron Acemoglu, a 2024 Nobel laureate in economics and Institute Professor of Economics at MIT, is a co-author (with Simon Johnson) of *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity* (PublicAffairs, 2023).

