DEMOCRACY AND SOCIETY 29.07.2026 | Kailash Satyarthi

Who does AI serve?

With AI, the moral questions matter first, because the answers will shape the future of humanity

The defining question of our time is no longer whether AI will transform our world. It already has. What the world needs now is a moral compass for the technology, which is why I introduced the concept of Compassionate AI to the 12 000-plus delegates from 170 countries who attended the AI for Good Global Summit in Geneva earlier this month.

If we are serious about ensuring that AI contributes to the common good, not just to profit and power, we must send AI engineers and developers out of their laboratories for one month. Let them meet children working in mines and factories, families displaced by conflict, and vulnerable communities enduring poverty, discrimination and exclusion. Let them walk beside people whose lives are shaped far more by uncertainty, fear, and injustice than by any technological advance. Even one month can change how a person thinks, and that will bear on how developers build this technology and shape the future.

The importance of compassion

AI is already reshaping education, health care, business, governance and scientific discovery. It promises remarkable advances, from accelerating medical research and improving learning to exposing trafficking networks, predicting hunger before famine strikes, and strengthening humanitarian action. But every technological revolution poses a moral test, and AI is no exception. With the technology having become the frontline of an intense global competition for economic and geopolitical power, speed is rewarded, scale is celebrated, and market dominance is prized above all else.

Under such conditions, we risk forgetting the most important question: Whom is this intelligence ultimately meant to serve? The greatest failure of AI today is one of leadership, not engineering. True, much of the conversation at global gatherings already focuses on concepts like ‘ethical AI’ and ‘responsible AI,’ and these are indispensable. We need stronger governance, meaningful accountability, and effective regulation.

But such safeguards are not enough. Ethics can tell us what is right, and laws can establish boundaries; but neither, by itself, will change the mindset of those who are imagining and designing technologies that will influence billions of lives. That requires something deeper, namely compassion.

Compassionate AI differs from every other framework because it begins not with technology, but with humanity.

Compassion is often mistaken for empathy, kindness, or charity. But it is not synonymous with any of them, nor is it simply an emotion or an abstract moral ideal. Compassion begins with recognising another person's suffering as our own, which prompts us to take action to alleviate that suffering. It includes the courage to act before additional harm occurs. It leads us not merely to optimise systems, but also to protect human dignity.

Far from being a constraint on technological progress, compassion is the backbone for responsible innovation. Compassionate AI differs from every other framework because it begins not with technology, but with humanity. Before asking what a system can do, Compassionate AI asks whom the system serves. Before pursuing greater efficiency, it considers the human consequences. Before celebrating innovation, it asks who might be left behind.

This single shift changes everything. Many of today's AI failures - bias, discrimination, exclusion, misinformation, environmental damage, threats to children's well-being - are often described as unfortunate side effects of progress. But they are not mere side effects. They are evidence that someone was absent, overlooked, or ignored when these systems were conceived.

Technology is not neutral, and neither are algorithms. Every system carries the assumptions, priorities, and values of the people who create it. Because AI was trained on humanity's accumulated knowledge, it has absorbed humanity's accumulated prejudices, inequalities and historical injustices.

Human rights cannot be the paint applied after the building is complete. They must be a load-bearing pillar of the entire structure. We trust that a bridge has been made safe before anyone crosses it, because safety is essential to its purpose. Obviously, AI should be held to the same standard. Compassion cannot be retrofitted after deployment. It must shape the system's design.

Changing the technology's trajectory is an urgent priority. Children growing up today may spend more time interacting with intelligent machines than with parents, teachers, or neighbours. We already understand some of the negative psychological consequences of social media, which is powered by relatively simple forms of AI. The effects of far more capable systems on children's emotional development, relationships, imagination, and mental well-being remain largely unknown, but they could be profound.

AI could become one of humanity's greatest allies.

Rather than dwelling primarily on how AI could transform the economy, we should be asking what will happen if algorithms replace the uniquely human capacity to feel another person's pain. Can they uphold justice if they were never designed to recognise suffering? Can they strengthen peace if they optimise only for engagement, efficiency, and profit? These moral questions must precede technical ones, because they could shape the future of humanity.

To be sure, AI could become one of humanity's greatest allies. It could help eradicate modern slavery and human trafficking, detect child sexual abuse, expand access to high-quality education in every child's own language, improve health care in underserved communities, predict humanitarian crises, and accelerate progress toward the Sustainable Development Goals.

But AI will not choose that future. We will. The principles we enshrine today will form the foundations for the increasingly powerful systems that we are building. Once those foundations harden into standards, institutions, and business models, they will shape humanity's relationship with intelligent machines for generations.

We cannot afford to wait until harms become irreversible. Legislation has often followed suffering; but with AI, we no longer have the luxury of a wait-and-see approach. The age of intelligence demands wisdom. History will not judge us by how intelligent our machines became. It will judge us by whether we had the foresight and courage to ensure that intelligence remained in the service of humanity.

© Project Syndicate



Kailash Satyarthi

Kailash Satyarthi is a Nobel Peace Prize laureate and Founder of Laureates and Leaders for Children, Honorary President of the Global March Against Child Labor, as well as the Founder of Bachpan Bachao Andolan (Save Childhood Movement).

-