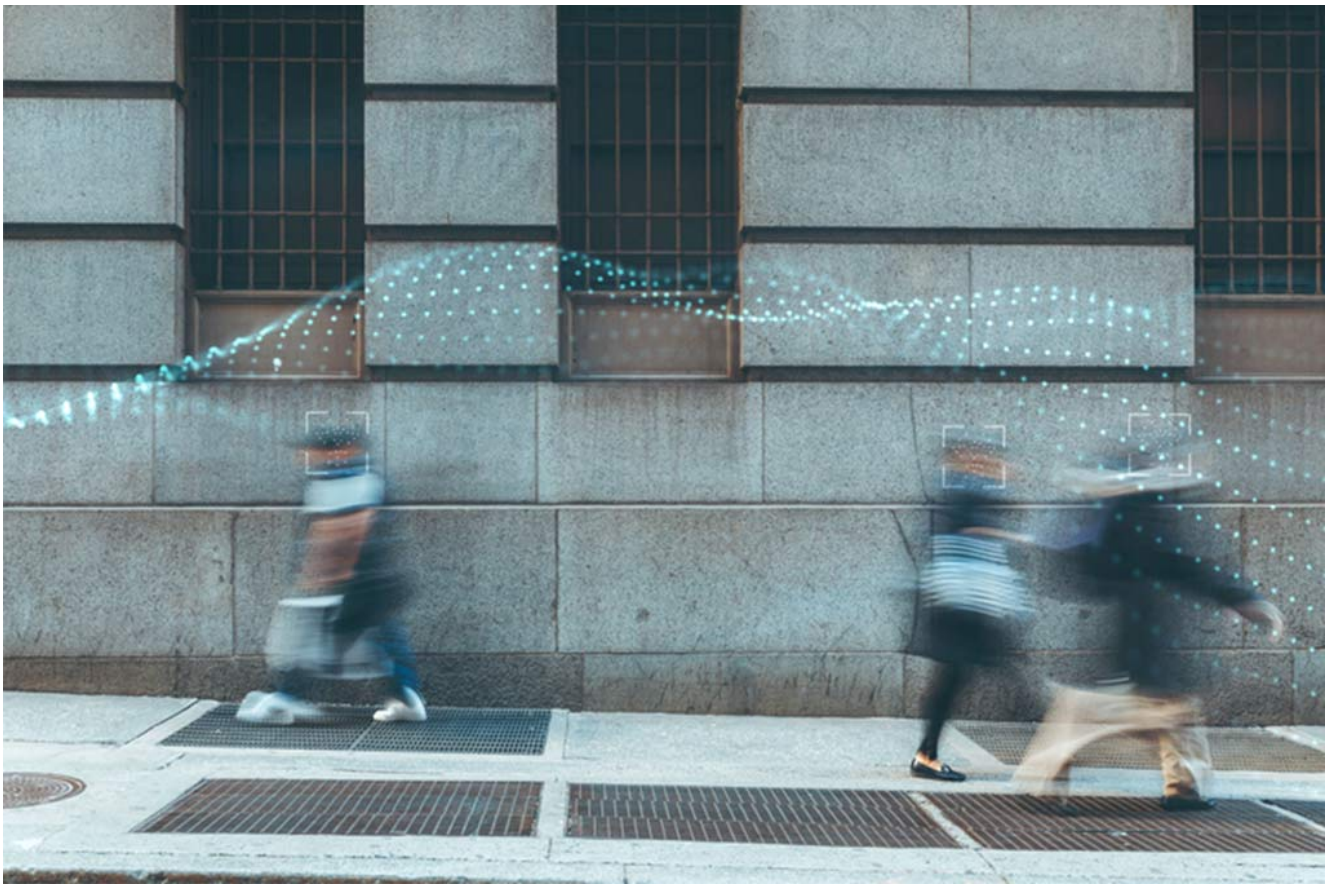
For AI development with ethical underpinnings, we should embrace humanism

With humanistic ethics, the starting point for AI projects is what the technology can do for humans or to them.



May 23, 2025

By Antonino Vaccaro and Rosa Fioravante

©IESE Business School

Much of the race to develop artificial intelligence (AI) is fueled by a belief that technological advancement is positive, in and of itself. It therefore fails to consider the human values embedded in technology.

This overreliance on tech's moral neutrality means that, rather than preventing ethical harms, organizations confront ethical issues after the fact, as unfortunate by-products of development. Companies rush to develop and commercialize AI, and then attempt to reduce sources of harm or correct models' failings afterward. This means they are consistently lagging behind in ethical compliance and oversight.

Humanistic ethics can help us to overcome this *ex-post* approach by offering a holistic and sound compass to organizations seeking to design and use people-centered technologies.

AI presents unprecedented ethical dilemmas arising from its potential to substitute human activities that were previously impossible to automate. Such dilemmas deepen with the specter of artificial general intelligence (AGI) or artificial super intelligence (ASI), with their ambitions to imitate and surpass the cognitive abilities of humans and act with autonomy. These technologies raise ethical concerns, not merely on occasional shortcomings such as bias or hallucinations, but on the boundaries of machine autonomy and the broader impact on humanity.

In our book, *Humanism and Artificial Intelligence*, which gathers contributions from more than a dozen global scholars across disciplines, we advocate for an *ex-ante* approach to AI ethical risks.

Instead of unfettered technological development being the default, we should begin by asking *why* a particular technology is being developed and *who* will benefit from its outputs.

To discuss the limits between what *can* be automated and what *ought* to remain under human decision-making, we deploy an applied ethics approach, aligned with that used in bioethics. In bioethics, fundamental questions around the ethical boundaries between humans and technological progress are at the core of the debate.

Humanistic ethics evaluates technological advancement through the lens of its human benefits, concerned first and foremost with the human purposes and needs to be met by innovation. Humanistic ethics considers *the human ends* of all economic activities and relationships transformed by AI, not simply the unethical by-products of certain technologies.

©IESE Business School

Humanism in AI in personal, organizational and societal dimensions

Humanism in business, and in particular the integral human development paradigm, can inform ethical AI in organizations on at least three pivotal levels:

- **Personal.** AI-led operations run the risk of prioritizing technological efficiency over employee wellbeing. Humanism and integral human development emphasize the interconnectedness of various dimensions of human life and advocate for policies and practices that enable individuals to flourish in all aspects of their lives. Organizations must avoid tendencies to devalue employees' experience through machine surveillance and control. In the case of employees involved in training AI, they must have access to specific upskilling and reskilling programs, to avoid being made redundant by the very AI they are helping to develop.
- **Organizational.** Humanism resists economic profitability, technological advancement and automation efficiency being the sole rationales of the organization, devaluing human moral values. It insists on the responsibility of organizations to promote conditions that enable individuals to achieve their full potential, ensuring spiritual and moral growth as well as material wellbeing. Corporations have underlying moral values that are evident in their cultures and structures. To transfer those moral values in AI adoption, organizations must build awareness of the technology across the company, involving all levels and departments to build trust toward both the technology itself and the governance decisions around it.
- **Societal.** Humanism informs stakeholder relationships, with a commitment to safeguard all people and groups, who possess intrinsic dignity and hold autonomous points of view. This understanding prevents them from being treated as mere means or passive subjects of automation processes or economic transactions. It promotes bringing stakeholders into AI decisions that potentially affect them. Reliance on machines over humans has been shown to facilitate unethical behavior in a variety of stakeholder relationships, as customers or others attempt to outwit technology, and vice versa. Involving critical stakeholders in technological disruption processes can smooth the transition and yield better outcomes for all parties.

©IESE Business School

AI and the risk of dehumanization

One of the reasons humanism is such a powerful antidote to uncritical technological optimism is because it helps to detect and prevent the risk of dehumanization, which may be involved in AI automation in at least two ways: by humanizing machines and by automating humans.

- In humanizing machines, organizations rely on AI's ability to perform human-like activities with increased efficiency, underestimating trade-offs arising from losing human components such as critical thinking in decision-making, empathy in diagnoses or knowledge of processes.
- In the mechanization of humans, people are treated as fungible in organizational dynamics and are thus subjected to the devices' work optimization, rather than subjecting technological advancements to human needs.

Humanistic ethics can help organizations to understand, analyze and face the risk of dehumanization. Rather than setting up ethical guardrails on technology that has already been released, it establishes the foundations on which artificial intelligence can be developed and deployed.

READ ALSO:

How to understand AI's potential impact on knowledge jobs

Read the future with science fiction and AI



Rosa Fioravante

Research fellow at IESE's Center for Business in Society and teaches a Business Ethics and Social Responsibility course at Católica Lisbon School of Business & Economics.



Antonino Vaccaro

Professor in the Business Ethics Department and the Negotiation Unit at IESE, and academic director of the Center for Business in Society. He specializes in compliance, social innovation and social entrepreneurship, and ethics and technology policy.

©IESE Business School

©IESE Business School