The Turbulent Era of AI and Artificial Wars: Gates' Warning and the Pentagon Clash

Published on 27/08/2026

Artificial intelligence represents **the greatest equalizer or the worst source of injustice** in human history. In an intervention published on *Gates Notes* on **August 26, 2026**, Bill Gates drew a clear and alarming balance sheet, defining the transition to this new era **as one of the most turbulent and unpredictable periods** for civilization. According to the founder of Microsoft, the world lacks **an adequate plan to govern a transformation** that compresses into a single decade epochal changes historically distributed over entire generations. Gates' words resonate as an urgent warning in the face of a reality knocking on the doors of geopolitics and global security, finding dramatic and immediate confirmation in recent events linking Silicon Valley giants to the U.S. Department of Defense.

In his essay, Gates articulates his concerns around **three fundamental pillars**: the need to create **new national and international institutions** dedicated to *AI governance*, the

deliberate protection of activities and professions through the concept of **Human Reserved**, and a **radical reform of taxation** that introduces targeted levies on artificial intelligence tokens and industrial robots. The stated goal is to discourage the indiscriminate replacement of human labor and finance flexible welfare. However, while the civil debate questions the employment impact — with research such as that from the **Stanford Digital Economy Lab** already highlighting an employment gap in young people between 22 and 25 years old in the most exposed professions — technological acceleration has found its most extreme battleground in the sector of defense and **algorithmic warfare**.

The convergence between frontier AI and military strategy: the OpenAI-Pentagon agreement

The convergence between frontier artificial intelligence and military planning represents **the most consequential and risky transformation** of modernity. The integration of intelligent systems into command and control networks raises profound questions about **strategic stability**, the preservation of human control over the use of lethal force, and the maintenance of **international humanitarian law**. This tension exploded dramatically with the agreement signed at the **end of February 2026** between OpenAI and the Pentagon, guided under the current administration by figures oriented toward muscular pragmatism in national security. The agreement, announced by **Sam Altman**, provides for the distribution of OpenAI's advanced models on classified infrastructures of the Department of Defense, accompanied by **rigorous formal safeguards**: absolute prohibition on domestic mass surveillance, total ban on fully autonomous lethal weapons systems, and maintenance of mandatory human oversight for every kinetic decision.

This development matured in the context of a deep fracture with **Anthropic** and its CEO **Dario Amodei**. Just hours before the OpenAI agreement, the U.S. administration had blacklisted Anthropic from federal agencies, even threatening to invoke the **Defense Production Act** and labeling the company as a "supply chain risk." The cause of the rupture lies in **Anthropic's categorical refusal** to remove the ethical guardrails that prevented the use of Claude models in internal surveillance activities and unsupervised targeting. This principled stance, while reinforcing the company's ethical image ahead of a historic stock market listing valued at **over**

three hundred billion dollars, highlighted the chasm separating the culture of civil research laboratories from the demands of geopolitical supremacy in an era of global conflicts, as evidenced by Middle Eastern tensions and fierce competition with China.

The structural dangers of algorithmic warfare and the human control crisis

The analysis of risks associated with military artificial intelligence, however, demonstrates how fragile the promises of containment based solely on corporate codes of conduct are. Literature on the dangers of military AI identifies **structural threats** that transcend individual commercial agreements. Among these, the development of **lethal autonomous weapons (LAWS)** and the progressive **erosion of human control** along a continuous spectrum represent the most visible danger. With the compression of decision-making times in the cyber and electronic domains, systems autonomously assume roles of threat identification and assessment, creating the phenomenon of **automation bias**, where the human operator blindly relies on algorithmic verdicts, transforming supervision into a formal ritual devoid of critical substance.

Even more insidious is the dynamic of **strategic stability and arms races**. The use of AI capable of identifying and striking strategic assets at a speed incomprehensible to the human mind erodes the survival of second-strike capabilities, encouraging scenarios of preventive first strike (**use-it-or-lose-it**). Unforeseen interactions between adversarial autonomous systems can trigger a chain of escalation at algorithmic speed, capable of transforming a local friction into a global conflict before any human actor can intervene to defuse it. To this is added the

fundamental problem of **algorithmic unpredictability**: the "black box" nature of neural networks and their intrinsic fragility in the face of adversarial examples or incomplete training data render combat systems vulnerable to **catastrophic errors** in real and complex operational contexts.

From legal responsibilities to the necessity of global governance

The legal and ethical implications invoke complex historical ghosts, including the principles established in the **Nuremberg trials** on individual and corporate responsibility of technology suppliers used for war crimes or humanitarian violations. When an artificial intelligence system fails to precisely distinguish between legitimate targets and civilians, or when the opacity of its calculations makes it impossible to establish who should answer penally for the error, **the entire architecture of international law of armed conflicts wavers**. Furthermore, the intrinsic vulnerability of global supply chains and machine learning software exposes defense networks to risks of tampering and advanced cyber attacks, capable of turning intelligent systems against their own creators.

The picture traced by Bill Gates and the reality of Pentagon military contracts converge toward one single, inescapable conclusion: the world finds itself **dangerously unprepared** in the face of the exponential speed of a technology that cuts across national security, economics, education, and social stability. Solutions cannot be entrusted to the benevolence or self-regulation of technology giants, whose commercial interests and competitive pressures constantly risk sacrificing ethics on the altar of strategic speed. As Gates warns, it is essential to establish **rigorous multilateral governance**, capable of coordinating States, experts, and civil society to preserve inviolable boundaries. If humanity fails in this task in the coming years, the turbulent era of AI will not limit itself to redefining the labor market or social relations, but will risk **handing the keys of life and death to machines devoid of conscience**, inaugurating a final chapter in which technology will cease to serve man to begin governing him.

News link: https://edition.cnn.com/2026/08/26/us/video/ac360gatesai